I.M. YAGLOM

COMPLEX NUMBERS IN GEOMETRY

Translator's Note

In the original book the references are all to Russian literature; wherever possible I have replaced these either by references to the corresponding English translations or to what, I hope, are equivalent references to books written in English and, occasionally, German.

A number of minor errors in the original edition have been corrected in this translation. Additions supplied by the author that could not be incorporated in the text have been placed at the end of the book, with references to them in the text.

I should like to acknowledge the great help which my wife has given in preparing the typescript.

Complex Numbers in Geometry

by I. M. YAGLOM
U.S.S.R.

Translated by ERIC J. F. PRIMROSE
UNIVERSITY OF LEICESTER
LEICESTER, ENGLAND

1968

ACADEMIC PRESS
New York and London

Preface

This book is intended for pupils in the top classes in high schools and for students in mathematics departments of universities and teachers' colleges. It may also be useful in the work of mathematical societies and may be of interest to teachers of mathematics in junior high and high schools.

The subject matter is concerned with both algebra and geometry. There are many useful connections between these two disciplines. Many applications of algebra to geometry and of geometry to algebra were known in antiquity; nearer to our time there appeared the important subject of analytical geometry, which led to algebraic geometry, a vast and rapidly developing science, concerned equally with algebra and geometry. Algebraic methods are now used in projective geometry, so that it is uncertain whether projective geometry should be called a branch of geometry or algebra. In the same way the study of complex numbers, which arises primarily within the bounds of algebra, proved to be very closely connected with geometry; this can be seen if only from the fact that geometers, perhaps, made a greater contribution to the development of the theory than algebraists.

Today different types of complex numbers are studied intensively; their study is connected with important unsolved problems, in which scientists from many countries are engaged. This book, naturally, does not aim to acquaint the reader with contemporary problems. Here only one of the threads that link the study of complex numbers with geometry will be shown and even in such a

limited field we shall not claim completeness. The range of questions touched on in this book, however, is sufficiently wide. In particular, we do not limit ourselves to the introduction of basic concepts, but in all cases we try to use these ideas to prove interesting geometrical theorems.

The book is intended for quite a wide circle of readers. The early sections of each chapter may be used in mathematical classes in secondary schools, and the later sections are obviously intended for more advanced students (this has necessitated a rather complicated system of notation to distinguish the various parts of the book).

The main line of the exposition is contained in Sections 1–4, 7, 9, 13, and 15, not denoted by asterisks. Parallel to this, and intended chiefly for present and future teachers of mathematics, there is a fairly wide selection of illustrations of an elementary geometrical character. The way to apply the apparatus of complex numbers to elementary geometry is demonstrated in Sections 8, 10, 14, and 16, which are denoted by *one asterisk.* Each of these four sections contains applications and examples and various geometrical theorems proved by using complex numbers. The theorems collected here are, as a rule, of purely illustrative significance; somewhat closer to the main line of the exposition are the theorems about the power of a point and a line with respect to a circle (Sections 8 and 10), which are used in Section 16 for a new geometrical definition of axial (Laguerre) inversion, which plays an essential role in the content of Section 15. The omission of Sections 8, 10, 14, and 16 will not affect the understanding of the rest of the book, and we recommend that the reader not spend very much time on them on a first reading; when he has mastered the basic material, the reader who is interested in elementary geometry can return to these sections.

Sections 5, 6, 11, 12, 17, and 18, denoted by *two asterisks*, are of a quite different character. Here we extend the bounds of the exposition and go beyond the material which, sometimes rather conventionally, is regarded as elementary geometry. The fact is that the chief application of complex numbers to geometry is still to Euclidean geometry, which is studied

in high schools, and to the so-called non-Euclidean geometries, of which the best known is that of Lobachevskii. Even in a popular book devoted to complex numbers, it seemed to us absolutely inadmissible to ignore completely this geometrical application of complex numbers. Although there is no possibility of dealing with this question in a short space, we have still thought it necessary to include in the book a short exposition of the role played by complex numbers in the geometry of Lobachevskii. The corresponding sections are naturally intended for the reader who has some familiarity with the content of this remarkable geometry. However, his preparation in this respect can be very little: it need not go beyond the limits of the material presented in popular science books and pamphlets on non-Euclidean geometry (some of these will be referred to in the footnotes). In accordance with the special character of the sections denoted by two asterisks, even the exposition there has a character rather different from the other parts of the book; for example, the proofs here are sometimes not completely carried out, and the filling in of some of the details is left to the reader. Thus, the omission of Sections 5, 6, 11, 12, 17, and 18 will not prevent the understanding of the rest of the book, which in its elementary parts (that is, those not connected with non-Euclidean geometry) forms a complete whole. We note that here and there in Sections 5, 6, 11, 12, 17, and 18 the account is somewhat more concise than that in the other parts of the book, and these sections contain hardly any concrete examples, similar to those contained in Sections 8, 10, 14 and 16, of the application of the techniques developed there; the independent construction of all the details of the proofs, and the carrying over to the non-Euclidean geometry of Lobachevskii of some of the results of Sections 8, 10, 14, and 16 may be regarded as problems which may properly be recommended to the student-reader.

The origin of this book was a lecture on the subject given by the author in 1958 to members of the school mathematical society attached to the Moscow state university. A broad account of this lecture was published in volume 6 of the collection *Mathe-*

matical Education (Fizmatgiz, Moscow) in 1961. A considerable part of the material was presented also to the society for students of the first course of the mathematical faculty of the Moscow state pedagogical institute. The author expresses his thanks to A. M. Yaglom, whose advice was taken during work on the manuscript, to his pupils D. B. Persits, M. M. Arapova, and F. M. Navyazhskii, who have provided some of the proofs carried out in the book, and to the editors of the book, M. M. Goryachaya and I. E. Morozova, who made a number of useful observations. Finally, he is grateful to R. Deaux, professor at the Polytechnic Institute of Mons (Belgium), who obligingly sent the latest edition of his book on complex numbers.

This English edition of the book differs from the original Russian edition in respect of a small number of corrections and additions, and also a brief Appendix, the basis of which is an article (in Russian) by the author: "Projective metrics in the plane and complex numbers" (*Proceedings of a Seminar on Vector and Tensor Analysis at Moscow State University*, Part 7, pp. 276–318, 1949).

The author thanks D. B. Persits for his constructive criticism of one part of the original Russian book; he is also grateful to the translator, Eric J. F. Primrose, for the great care with which he has carried out the work of preparing the English edition, and to Academic Press for their considerate treatment of all the author's requests.

I. M. YAGLOM

Contents

Chapter III: *Circular Transformations and Circular Geometry*

Appendix: *Non-Euclidean Geometries in the Plane and Complex Numbers*

CHAPTER I

Three Types of Complex Numbers

§1. Ordinary Complex Numbers

The introduction of complex numbers into algebra is connected with the solution of quadratic equations. If by "numbers" we understand only the usual real numbers, then we must say that the quadratic equation

$$x^2 + px + q = 0 \tag{1}$$

has two roots if $\Delta = p^2 - 4q > 0$, one root if $\Delta = 0$, and no root if $\Delta < 0$. Thus, very many equations, such as

$$x^2 + 1 = 0, \qquad x^2 - 2x + 2 = 0, \qquad x^2 + x + 1 = 0 \tag{2}$$

prove to be insoluble, because they have no roots. This situation considerably complicates the theory of equations.

To eliminate this complication it is convenient to *extend the idea of number*. That is, we agree to say that the equation $x^2 + 1 = 0$ has a root, which is *a number of a special kind*, different from the usual real numbers. It is called **imaginary** and is denoted by the special letter i. If we adjoin the number i to the set of real numbers, we must explain how to multiply real numbers by i and how to add them to i; as we know, we can multiply numbers together and add them together, and until we define these operations for our extended set of numbers we have no sufficient basis for calling i a number. Here it turns out to be impossible to confine ourselves to the adjunction of only one number i; all products bi of a real number b by i and all

sums $a + bi$ of a real number a and a number bi (where $b \neq 0$) must also count as numbers of a special kind and be included in the new set with the real numbers and the number i. The set of numbers of the form $a + bi$ obtained in this way ($b = 0$ gives all real numbers and $a = 0$ gives numbers of the form bi) is called the set of **complex numbers**.

Addition, subtraction, and multiplication of complex numbers are naturally defined in the following way:

$$\begin{aligned}(a + bi) + (c + di) &= (a + c) + (b + d)i \\ (a + bi) - (c + di) &= (a - c) + (b - d)i \\ (a + bi)(c + di) &= ac + adi + bci + bdi^2 \\ &= (ac - bd) + (ad + bc)i\end{aligned} \tag{3}$$

Here we use the fact that *by definition* i is a root of the equation $x^2 + 1 = 0$, so $i^2 + 1 = 0$ and $i^2 = -1$.

It is just as easy to explain the rule for the division of a complex number by a real number:

$$\frac{(c + di)}{a} = (c + di)\frac{1}{a} = \frac{c}{a} + \frac{d}{a}i$$

If we need to divide an arbitrary complex number z_1 by another complex number z, it is sufficient to choose a number $\bar{z}$, such that the product $z\bar{z}$ is *real*. Then we shall have

$$z_1/z = z_1\bar{z}/z\bar{z} \tag{4}$$

and the rules for multiplying the complex numbers z_1 and $\bar{z}$ and dividing the resulting complex number $z_1\bar{z}$ by the real number $z\bar{z}$ are already known. Let $z = a + bi$; in this case it is convenient to choose for $\bar{z}$ the number $a - bi$, for which

$$z\bar{z} = (a + bi)(a - bi) = a^2 + b^2 \tag{5}$$

Now the rule (Equation 4) for dividing by the complex number $z = a + bi$ can be written as

$$\begin{aligned}\frac{c + di}{a + bi} = \frac{(c + di)(a - bi)}{(a + bi)(a - bi)} &= \frac{(ca + db) + (-cb + da)i}{a^2 + b^2} \\ &= \frac{ca + db}{a^2 + b^2} + \frac{-cb + da}{a^2 + b^2}i\end{aligned} \tag{6}$$

The number $\bar{z} = a - bi$ is called **conjugate**[1] to the complex number $z = a + bi$; it is obvious that, conversely, z is conjugate to $\bar{z}$ (i.e. $\overline{(\bar{z})} = z$). We observe that not only the product $z\bar{z}$, but also the sum $z + \bar{z}$, of conjugate complex numbers is a real number. The sum $z + \bar{z} = 2a$ is twice the *real part* a of the complex number $z = a + bi$; the product $z\bar{z} = a^2 + b^2$ is the square of the (positive) number $r = +(a^2 + b^2)^{1/2}$, which is called the **modulus** of z and is denoted by $|z|$. It is also obvious that a number z is equal to its conjugate (that is, $\bar{z} = z$), if and only if z is a *real* number. Further, it is easy to show that from the definition of a conjugate number follow the identities

$$\bar{z}_1 + \bar{z}_2 = \overline{z_1 + z_2}\,, \qquad \bar{z}_1 - \bar{z}_2 = \overline{z_1 - z_2}$$
$$\bar{z}_1 \cdot \bar{z}_2 = \overline{z_1 \cdot z_2}\,, \qquad \bar{z}_1/\bar{z}_2 = \overline{z_1/z_2} \tag{7}$$

Put into words, the sum, difference, product, and quotient of the numbers conjugate to two given complex numbers are respectively conjugate to the sum, difference, product, and quotient of those numbers. Hereafter we shall need also the fact that the difference $z - \bar{z}$ of two conjugate numbers is a **purely imaginary** number (that is, it has the form bi, where b is real).[2]

So complex numbers can be added, subtracted, multiplied, and divided, and all the laws which these operations satisfy agree with the laws of operation for ordinary real numbers.[3] In particular, as in the case of real numbers, division by a complex number $z = a + bi$ is *not always* possible; for division to be possible it is essential that the modulus $|z| = (a^2 + b^2)^{1/2}$ of z shall be nonzero. Thus, there exists a unique complex number $0 = 0 + 0i$, for which division is impossible. In those cases in which the impossibility of dividing by zero is inconvenient we proceed in the usual way: we agree to say that the quotient $1/0$ exists, but that it is *a number of a special kind*, for which we

---

[1] Hereinafter the symbol $\bar{z}$ will always denote the number conjugate to z.

[2] If $z = c + di$, then, of course, $z - \bar{z} = 2di$.—Transl.

[3] This property is sometimes expressed by saying that complex numbers, just as real numbers, form a *commutative field.*

introduce the special symbol ∞. In other words, we extend the set of complex numbers by introducing the new number ∞, which by definition is equal to $1/0$. The rules for operating with the symbol ∞ are defined in the following way:

$$z + \infty = \infty, \qquad z - \infty = \infty,$$
$$z \cdot \infty = \infty, \qquad \frac{\infty}{z} = \infty, \qquad \frac{z}{\infty} = 0 \tag{8}$$

Here z is an arbitrary number; in the third equation $z \neq 0$, and in the second and last two equations $z \neq \infty$. The difference $\infty - \infty$, the product $0 \cdot \infty$, and the ratio ∞/∞ and, we may add, also the ratio $0/0 = 0 \cdot (1/0) = 0 \cdot \infty$, must be regarded, generally speaking, as having no meaning, and nothing can be done about them.[4]

It is important to keep in mind the fact that, whereas in the field of real numbers the extraction of a square root is possible only in the case of a positive number (more precisely, of a non-negative number), in the field of complex numbers the square root of *any* number $z = a + bi$ can be extracted. In fact, if we put

$$a + bi = (x + yi)^2$$

we easily obtain

$$x^2 - y^2 = a, \qquad 2xy = b$$

Solving these equations, we obtain $x = [(|z| + a)/2]^{1/2}$ and $y = [(|z| - a)/2]^{1/2}$, where $z = (a^2 + b^2)^{1/2} \geqslant a$ (the signs of the square roots for x and y are chosen so that the product xy has the same sign as b). This leads to the formulae

$$(a + bi)^{1/2} = \pm \left[\left(\frac{|z| + a}{2} \right)^{1/2} + \left(\frac{|z| - a}{2} \right)^{1/2} i \right]$$
$$(a - bi)^{1/2} = \pm \left[\left(\frac{|z| + a}{2} \right)^{1/2} - \left(\frac{|z| - a}{2} \right)^{1/2} i \right], \quad b \geqslant 0 \tag{9}$$

[4] We note, however, that the ratio $(a \cdot \infty + b)/(c \cdot \infty + d)$, where a, b, c, and d are arbitrary complex numbers, should be reckoned, by means of the identities $(az + b)/(cz + d) \equiv [a + b(1/z)]/[c + d(1/z)]$ and $1/\infty = 0$, as having a quite definite meaning, namely a/c. This remark will be useful to us later.

Hence it follows easily that *in the domain of complex numbers every quadratic equation* (Equation 1, with real or complex coefficients p and q) *has two* (distinct or coincident) *roots*:

$$x_{1,2} = \frac{-p \pm (p^2 - 4q)^{1/2}}{2} \tag{10}$$

In particular, for *real* p and q this equation will have two distinct real roots $x_{1,2} = (-p \pm \Delta^{1/2})/2$ if $\Delta > 0$, two coincident (real) roots $x_{1,2} = -p/2$ if $\Delta = 0$, and two distinct complex (mutually conjugate) roots $x_{1,2} = (-p \pm (-\Delta)^{1/2} i)/2$ if $\Delta < 0$. So, for example, Equations 2 have the following roots:

$$x_{1,2} = \pm i, \qquad x_{1,2} = 1 \pm i, \qquad x_{1,2} = -\frac{1}{2} \pm \frac{\sqrt{3}}{2} i \tag{11}$$

In many cases it is more convenient to write a complex number $z = a + bi$ in a different form, by emphasizing its modulus $|z| = (a^2 + b^2)^{1/2}$:

$$z = a + bi = (a^2 + b^2)^{1/2} \left(\frac{a}{(a^2 + b^2)^{1/2}} + \frac{b}{(a^2 + b^2)^{1/2}} i \right)$$

The real numbers $a/(a^2 + b^2)^{1/2}$ and $b/(a^2 + b^2)^{1/2}$ have the property that the sum of their squares is equal to 1. Hence follows the existence of an angle φ, such that

$$\cos \varphi = \frac{a}{(a^2 + b^2)^{1/2}} \qquad \sin \varphi = \frac{b}{(a^2 + b^2)^{1/2}} \tag{12}$$

If the modulus $|z| = (a^2 + b^2)^{1/2}$ of the number z is denoted by the letter r, we have

$$z = r(\cos \varphi + i \sin \varphi) \tag{13}$$

where

$$r = (a^2 + b^2)^{1/2}, \quad \cos \varphi = \frac{a}{r}, \quad \sin \varphi = \frac{b}{r} \tag{13a}$$

The angle φ (defined by Equations 13a, apart from the addition of any multiple of 2π) is called the **argument** of z and is denoted

by arg z; if it is restricted, say by the condition $-\pi < \varphi \leqslant \pi$, then for positive real numbers it is equal to 0, and for negative real numbers it is equal to π. Conjugate numbers will have the same modulus r and opposite arguments φ and $-\varphi$.

The form of writing complex numbers shown in Equation 13 is called the **trigonometrical form**. It is extremely convenient when two or more complex numbers must be multiplied. In fact,

$$\begin{aligned} &r(\cos\varphi + i\sin\varphi)\cdot r_1(\cos\varphi_1 + i\sin\varphi_1)\\ &= rr_1[(\cos\varphi\cos\varphi_1 - \sin\varphi\sin\varphi_1) + i(\cos\varphi\sin\varphi_1 + \sin\varphi\cos\varphi_1)]\\ &= rr_1[(\cos(\varphi+\varphi_1) + i\sin(\varphi+\varphi_1)] \end{aligned} \tag{14}$$

Thus, *the modulus of the product of two complex numbers is equal to the product of the moduli of the factors, and the argument of the product is equal to the sum of the arguments of the factors* (cf. the much less convenient Equation 3). Hence it follows that *the modulus of the quotient of two complex numbers is equal to the quotient of the moduli of the numbers, and the argument of the quotient is equal to the difference of the corresponding arguments*:

$$\frac{z_1}{z} = \frac{r_1(\cos\varphi_1 + i\sin\varphi_1)}{r(\cos\varphi + i\sin\varphi)} = \frac{r_1}{r}[\cos(\varphi_1 - \varphi) + i\sin(\varphi_1 - \varphi)] \tag{15}$$

From these rules certain laws, which show how to raise a complex number to any power and to extract any root of it, follow immediately:

$$\begin{aligned} [r(\cos\varphi + i\sin\varphi)]^n &= r^n(\cos n\varphi + i\sin n\varphi)\\ [r(\cos\varphi + i\sin\varphi)]^{1/n} &= r^{1/n}\left(\cos\frac{\varphi}{n} + i\sin\frac{\varphi}{n}\right) \end{aligned} \tag{16}$$

We obtain the n different values of the nth root by taking for φ in the latter formula the n values of the argument $\varphi = \varphi_0 + 2k\pi$, where $k = 0, 1, \ldots, n-1$ and φ_0 is any one of the possible values of the argument.

It is interesting to note that the purely formal method of introducing complex numbers mentioned here is very general and can be used at

the very beginning of a course in algebra to introduce rational (fractional) numbers and directed (positive and nonpositive) numbers. In fact, if we have only positive integers (and zero) we can easily add and multiply them, but subtraction is not always possible. In order that subtraction may be possible (so as to obtain solutions of all equations $x + a = b$), it is necessary to extend the set of positive numbers by adjoining as a "number of a special kind" the root of the equation $x + 1 = 0$, which is denoted by -1; further, by addition and multiplication all negative integers can be obtained, and so all equations $x + a = b$ (with integral coefficients) have solutions. In the same way, in order that division may be possible (so as to obtain solutions of all equations $ax = b$ with integral coefficients), it is necessary to extend the set of numbers still further, by introducing as "numbers of a special kind" the solutions $1/a$ of all equations $ax = 1$, where a is a positive integer; later we naturally proceed to the set of fractions (rational numbers) b/a, and then all equations of the first degree with integral coefficients, $ax = b$, become soluble. In the same way we may introduce quadratic irrationals and irrationals of higher degree.

We emphasize that the fundamental importance of complex numbers for algebra stems primarily from the fact that in going over from the quadratic Equation 1 to equations of higher degree it is not necessary to extend the set of numbers further by adjoining to numbers of the form $a + bi$ any more "numbers of a special kind"; it turns out that *any equation of degree n with real or arbitrary complex coefficients always has a complex root.* This is known as the **fundamental theorem of algebra.**

§2. Generalized Complex Numbers

Let us return to the beginning of the path which leads to the construction of complex numbers. To remove the difficulties connected with the insolubility of some quadratic equations in the field of real numbers we adjoined to the set of real numbers a new element i, which *by definition* is a root of one of the insoluble equations, namely $x^2 + 1 = 0$; this led to the set of complex numbers $a + bi$, where a and b are real, in terms of which, as we have shown, all quadratic equations have roots. Let us now ask the question whether it is essential in this construction to use the equation $x^2 + 1 = 0$, or whether it might be replaced with some other quadratic equation?

The answer to this question is not difficult. It is easily seen that the equation $x^2 + 1 = 0$ has *no advantage*, in principle, over other equations insoluble in the real field, and the choice of it is dictated only by its relative simplicity (that is, the coefficients p and q are equal to 0 and 1). In fact, let us denote by I a "number of a special kind," which is by definition a root of an arbitrary quadratic equation with negative discriminant,

$$x^2 + px + q = 0, \qquad \Delta = p^2 - 4q < 0 \tag{17}$$

and let us consider the set of **generalized complex numbers** Z of the form

$$a + bI, \qquad a, b \text{ real} \tag{18}$$

These numbers may be added, subtracted, and multiplied according to the rules

$$\begin{aligned}(a + bI) + (c + dI) &= (a + c) + (b + d)I \\ (a + bI) - (c + dI) &= (a - c) + (b - d)I \\ (a + bI)(c + dI) &= ac + adI + bcI + bdI^2 \\ &= (ac - qbd) + (ad + bc - pbd)I\end{aligned} \tag{19}$$

$I^2 = -pI - q$, since I is *by definition* a root of the equation $x^2 + px + q = 0$. Now, it is easy to choose for each generalized complex number $Z = a + bI$ a number $\bar{Z}$ such that the product $Z\bar{Z}$ is real. For example, we can put $\bar{Z} = (a - pb) - bI$; then

$$Z\bar{Z} = a^2 - pab + qb^2 = \left(a - \frac{p}{2}b\right)^2 + \frac{4q - p^2}{4}b^2$$

This allows us to define division of generalized complex numbers, as in Equation 4; since, moreover, $Z\bar{Z} = 0$ only if $a = 0$ and $b = 0$ (because $(4q - p^2)/4 = -\Delta/4 > 0$), the only number by which it is impossible to divide is 0 (which is equal to $0 + 0 \cdot I$). Finally, it is easy to show that every quadratic equation (with real or generalized complex coefficients) has in the field of generalized complex numbers two (coincident or distinct) roots.

So, for example, if I denotes a root of the second of Equations 2, the roots of the three equations are:

$$x_1 = -1 + I, \qquad x_2 = 1 - I$$

$$x_1 = I, \qquad x_2 = 2 - I$$

$$x_1 = \frac{\sqrt{3} - 1}{2} - \frac{\sqrt{3}}{2} I, \qquad x_2 = -\frac{\sqrt{3} + 1}{2} + \frac{\sqrt{3}}{2} I$$

If I_1 is a root of the third of Equations 2, then the roots of the equations have the form

$$x_1 = \frac{\sqrt{3}}{3} + \frac{2\sqrt{3}}{3} I_1, \qquad x_2 = -\frac{\sqrt{3}}{3} - \frac{2\sqrt{3}}{3} I_1$$

$$x_1 = \frac{3 + \sqrt{3}}{3} + \frac{2\sqrt{3}}{3} I_1, \qquad x_2 = \frac{3 - \sqrt{3}}{3} - \frac{2\sqrt{3}}{3} I_1$$

$$x_1 = I_1, \qquad x_2 = -1 - I_1$$

All these results become quite obvious when we remember that the root I of Equation 17 has the form

$$I = -\frac{p}{2} + \frac{(-\Delta)^{1/2}}{2} i, \qquad \text{or} \quad I = -\frac{p}{2} - \frac{(-\Delta)^{1/2}}{2} i \tag{20}$$

and, conversely, i can be expressed in terms of I:

$$i = \frac{p}{(-\Delta)^{1/2}} + \frac{2}{(-\Delta)^{1/2}} I, \qquad \text{or} \quad i = -\frac{p}{(-\Delta)^{1/2}} - \frac{2}{(-\Delta)^{1/2}} I \tag{21}$$

Thus, *generalized complex numbers $a + bI$ are the same as ordinary complex numbers $a + bi$ but are expressed in a somewhat different form*:

$$a + bI = a + b\left(-\frac{p}{2} + \frac{(-\Delta)^{1/2}}{2} i\right) = a_1 + b_1 i$$

where

$$a_1 = a - \frac{p}{2} b, \qquad b_1 = \frac{(-\Delta)^{1/2}}{2} b \tag{22}$$

Thus it is clear that all the algebraic properties of the numbers

$Z = a + bI$ are no different from the properties of ordinary complex numbers.[5]

§3. The Most General Complex Numbers

We now take one more step in the direction of a further generalization of the idea of a complex number. We set ourselves the problem of how far it is essential in the developments given in the preceding section that the discriminant Δ of Equation 17 be negative and whether it is impossible to get rid of this restriction.

It is clear that if we look at complex numbers as a means of combatting the difficulties connected with the insolubility of some quadratic equations in the domain of real numbers, then the condition that the discriminant Δ be negative is absolutely essential. In fact, let us suppose that we adjoin to the set of ordinary real numbers a number E as a "number of a special kind," which is by definition a root of Equation 1 with *positive* discriminant. In this case, in the domain of numbers of the form $a + bE$ Equation 1 will have at least three different roots: two real roots, given by Equation 10 (since $\Delta = p^2 - 4q > 0$), and the root E (in fact, there will be four different roots); at the same time it is not difficult to show that all equations with negative discriminant remain insoluble.[6] Henceforth, however, we shall no longer consider the question of the solubility of quadratic equations,

---

[5] The identification of the algebraic properties of generalized complex numbers Z and ordinary complex numbers z follows from the *isomorphism* of these two sets of numbers, that is, from the existence of a one-to-one mapping $z \leftrightarrow Z$ such that, if $z_1 \leftrightarrow Z_1$ and $z_2 \leftrightarrow Z_2$, then $z_1 + z_2 \leftrightarrow Z_1 + Z_2$ and $z_1 - z_2 \leftrightarrow Z_1 - Z_2$ and $z_1 \cdot z_2 \leftrightarrow Z_1 \cdot Z_2$ and $z_1/z_2 \leftrightarrow Z_1/Z_2$ (we obtain this mapping by making the number $z = a + bi$ correspond to the number $Z = a_1 + b_1 I$, where $a_1 = a + pb/(-\Delta)^{1/2}$ and $b_1 = 2b/(-\Delta)^{1/2}$ and, hence, $a = a_1 - (p/2)b_1$ and $b = \frac{1}{2}(-\Delta)^{1/2} b_1$).

[6] We give the corresponding proof for the special case of an equation of the form $x^2 + c = 0$, where $c > 0$. From the fact that $E^2 = -pE - q$ it turns out that $(a + bE)^2 = (a^2 - qb^2) + (2ab - pb^2)E$; therefore $(a + bE)^2$ is a real number only when $2ab - pb^2 = 0$, that is, when $b = 0$ or $b = (2/p)a$. But then either $(a + bE)^2 = a^2 > 0$, if $b = 0$,

but we shall consider complex numbers only as certain numbers of a new kind, related to ordinary numbers by rules for the application of algebraic operations and having (as we shall see below) interesting applications to geometry.

From this new point of view the extension of the set of real numbers by adjoining a new element E, satisfying Equation 1 *by definition*, is legitimate, regardless of the sign of the discriminant Δ of this equation. All possible linear combinations

$$a + bE, \qquad a, b \text{ real} \tag{23}$$

will be called the **most general complex numbers**.

Addition, subtraction, and multiplication of the most general complex numbers will be carried out according to the following natural laws:

$$\begin{aligned}(a + bE) + (c + dE) &= (a + c) + (b + d)E \\ (a + bE) - (c + dE) &= (a - c) + (b - d)E \\ (a + bE) \cdot (c + dE) &= ac + adE + bcE + bdE^2 \\ &= (ac - qbd) + (ad + bc - pbd)E\end{aligned} \tag{24}$$

It is not difficult to see that all the laws concerning addition, subtraction, and multiplication of the most general complex numbers agree with the laws of operation for ordinary real numbers. The question of division of these numbers is somewhat different (in the analysis of the question of division of generalized complex numbers in the preceding section we used the fact that the discriminant $\Delta = p^2 - 4q$ was negative); hence, for the present we shall leave open this question.[7]

---

or $(a + bE)^2 = a^2(1 - 4q/p^2) > 0$, if $b = (2/p)a$; here we use the fact that, by hypothesis, $p^2 - 4q > 0$. Thus, for a number $x = a + bE$ of this kind, x^2 *cannot* be equal to the negative number $-c$. Similarly, we may show that an *arbitrary* quadratic equation with negative discriminant is insoluble in the domain of numbers of the form $a + bE$.

[7] The fact that the most general complex numbers can be added, subtracted, and multiplied, all the usual laws of these operations being conserved, but that it is not always possible to divide one by another is expressed by saying that such numbers form a **ring**.

There are very many systems of the most general complex numbers. Any pair of real numbers p and q can correspond to the quadratic Equation 1 and, consequently, to a system of the most general complex numbers, Equation 23. However, as we already know from the previous section, all such systems of numbers corresponding to pairs p and q such that

$$\Delta = p^2 - 4q < 0$$

do not essentially differ from each other; among numbers of the form $a + bE$ there is always a number $i = \alpha + \beta E$ such that $i^2 = -1$ (for this it is only necessary to put $\alpha = p/(-\Delta)^{1/2}$ and $\beta = 2/(-\Delta)^{1/2}$, and then $E = \alpha_1 + \beta_1 i$, where $\alpha_1 = -p/2$ and $\beta_1 = (-\Delta)^{1/2}/2$, and therefore the number $a + bE$ can be identified with the ordinary complex number $a_1 + b_1 i$, where $a_1 = a - pb/2$ and $b_1 = \frac{1}{2}(-\Delta)^{1/2}b$). Similarly, *in the case in which*

$$\Delta = p^2 - 4q = 0,$$

among numbers of the form $a + bE$ *there is a number* $\varepsilon = \alpha + \beta E$ *such that* $\varepsilon^2 = 0$; it is possible, for example, to put $\varepsilon = p/2 + E$; then

$$\varepsilon^2 = \frac{p^2}{4} + pE + (-pE - q) = \frac{p^2}{4} - q = 0$$

Hence, the aggregate of the most general complex numbers $a + bE$, where $p^2 - 4q = 0$, can always be reduced to the so-called **dual numbers**:

$$a + b\varepsilon, \qquad a, b \text{ real}, \quad \varepsilon^2 = 0 \tag{25}$$

The number $a + bE$ must be identified with the dual number $a_1 + b_1$, where $a_1 = a - (p/2)b$ and $b_1 = b$.

Finally, *if we have*

$$\Delta = p^2 - 4q > 0$$

then there are complex numbers $e = a + bE$ *such that* $e^2 = +1$; in fact, if we put

$$e = \frac{p}{\Delta^{1/2}} + \frac{2}{\Delta^{1/2}} E, \qquad E = -\frac{p}{2} + \frac{\Delta^{1/2}}{2} e$$

then we have

$$e^2 = \left(\frac{p}{\Delta^{1/2}} + \frac{2}{\Delta^{1/2}} E\right)^2 = \frac{p^2}{\Delta} + \frac{4p}{\Delta} E + \frac{4}{\Delta}(-pE - q) = 1$$

This allows us to reduce our system of most general complex numbers to the so-called **double numbers**:

$$a + be, \qquad a, b \text{ real}, \quad e^2 = 1 \tag{26}$$

It is sufficient to identify the number $a + bE$ with the double number

$$a_1 + b_1 e = \left(a - b\frac{p}{2}\right) + \left(b\frac{\Delta^{1/2}}{2}\right) e$$

So we see that *all systems of the most general complex numbers can, in fact, be reduced to the following three different systems*[8]:

Ordinary complex numbers	$a + bi, \quad i^2 = -1$
Dual numbers	$a + b\varepsilon, \quad \varepsilon^2 = 0$
Double numbers	$a + be, \quad e^2 = 1$

Ordinary complex numbers are closely connected with the question of the solution of equations of the second degree and higher; they play a fundamental role in algebra and in many parts of mathematical analysis. Their origin is difficult to trace. It is believed that the first to make use of them were the sixteenth-century Italian mathematicians G. Cardan (1501–1576) and R. Bombelli (born in 1530, his *Algebra* published in 1572),

[8] More precisely, the most general complex numbers are isomorphic with ordinary complex numbers when $\Delta < 0$, with dual numbers when $\Delta = 0$, and with double numbers when $\Delta > 0$ (cf. footnote 5). These numbers are sometimes called *elliptic complex numbers* when $\Delta < 0$, *parabolic complex numbers* when $\Delta = 0$, and *hyperbolic complex numbers* when $\Delta > 0$.

TRANSLATOR'S NOTE: There seems to be some disagreement in the literature about these terms. H. Busemann in a review (*Mathematical Reviews*, Vol. 12, p. 351) of a paper by Yaglom uses the term *elliptic complex number* for "double number." Not having then read the present book, I used this term when translating the paper referred to in footnote 11.

but these numbers can be found implicit in much earlier works; on the other hand, long after the times of Cardan and Bombelli even eminent mathematicians did not have a proper idea about complex numbers. Dual and double numbers do not have any connection with the theory of quadratic equations with real coefficients and in general have relatively little connection with algebra; their main application is in geometry.[9] Dual numbers, apparently, were first considered by the famous German geometer E. Study (1862–1930) of the end of the last century and the beginning of this one; double numbers were introduced by a contemporary of Study, the English geometer W. Clifford (1845–1879).[10]

The chief application of double numbers is concerned with the non-Euclidean geometry of Lobachevskii[11]; hence, in this book we concentrate our attention first on ordinary complex numbers and dual numbers.

§4. Dual Numbers

Addition, subtraction, and multiplication of dual numbers are defined by the formulae

$$\begin{aligned}(a + b\varepsilon) + (c + d\varepsilon) &= (a + c) + (b + d)\varepsilon \\ (a + b\varepsilon) - (c + d\varepsilon) &= (a - c) + (b - d)\varepsilon \\ (a + b\varepsilon)(c + d\varepsilon) &= ac + (ad + bc)\varepsilon\end{aligned} \tag{27}$$

[9] Some applications of these systems of complex numbers are made in the theory of numbers.

[10] Clifford, who was concerned with the use of these numbers in mechanics, called them *motors*.—TRANSL.

[11] Double numbers enter also into other geometries different from the usual geometry of Euclid (for example, the so-called **pseudo-Euclidean geometry**, which plays a fundamental role in the theory of relativity; see Appendix). [The question of the connection of complex numbers with non-Euclidean geometries is investigated in the paper "Projective metrics," by I. M. Yaglom, B. A. Rozenfel'd, and E. U. Yasinskaya, in *Russian Mathematical Surveys*, Vol. 19, No. 5 (1964), pp. 49–107; the reader should be warned that this paper is by no means elementary and is intended for trained teachers.]

The last of these formulae shows that the product of a dual number $z = a + b\varepsilon$ and another number $z_1 = c + d\varepsilon$ is particularly simple in the case in which $ad + bc = 0$; if $a \neq 0$, this last equation can be written in the form $b/a = -d/c$. In particular, the product of two numbers $z = a + b\varepsilon$ and $\bar{z} = a - b\varepsilon$ is real:

$$z \cdot \bar{z} = (a + b\varepsilon)(a - b\varepsilon) = a^2 \tag{28}$$

The number $\bar{z} = a - b\varepsilon$ is called **conjugate** to the number $z = a + b\varepsilon$ (and, conversely, z is conjugate to $\bar{z}$); the square root a of the product $z\bar{z}$ (which coincides with half the sum, $(z + \bar{z})/2$, of the conjugate numbers z and $\bar{z}$) is called the **modulus** of the dual number z and is denoted by $| z |$ (we note that the modulus of a dual number can even be negative!). The sum $z + \bar{z} = 2a$ of two conjugate numbers is real; the difference $z - \bar{z} = 2b\varepsilon$ is a purely imaginary number (that is, it differs from ε only by a real multiplier). We notice that, just as in ordinary complex numbers, a dual number z coincides with its conjugate $\bar{z}$ if and only if it is real. Moreover, Equations 7 all remain valid for dual numbers.

The rule for division by a dual number $z = a + b\varepsilon$ may now be expressed as

$$\begin{aligned}\frac{c + d\varepsilon}{a + b\varepsilon} &= \frac{(c + d\varepsilon)(a - b\varepsilon)}{(a + b\varepsilon)(a - b\varepsilon)} = \frac{ca + (-cb + da)\varepsilon}{a^2} \\ &= \frac{c}{a} + \frac{-cb + da}{a^2}\,\varepsilon\end{aligned} \tag{29}$$

Hence it is clear that for division by a dual number z to be possible it is essential that the modulus $| z | = a$ of the number be *nonzero*; here, in contrast to ordinary complex numbers, a dual number with zero modulus can be nonzero. In those cases in which the impossibility of dividing by a number with zero modulus causes us difficulty we may regard the quotients $1/\varepsilon$ and $1/0$ as numbers of a different character, which we agree to denote by ω and ∞; let us also introduce all possible numbers

of the form $c\omega$, where $c \neq 0$ is real. Then any dual number will have an inverse:

$$\frac{1}{b\varepsilon} = \frac{1}{b}\,\omega, \quad \text{if } b \neq 0; \qquad \frac{1}{0} = \infty$$

The rule for division by the symbol ∞ is defined here by Equations 8 (where the number z can be a number of the form $c\omega$); the rules of operation with numbers $c\omega$ are defined thus[12]:

$$\begin{gathered}(a + b\varepsilon) + c\omega = c\omega, \qquad (a + b\varepsilon) - c\omega = (-c)\omega \\ (a + b\varepsilon)c\omega = (ac)\omega \\ \frac{c\omega}{a + b\varepsilon} = \frac{c}{a}\,\omega, \qquad \frac{a + b\varepsilon}{c\omega} = \frac{a}{c}\,\varepsilon \\ c\omega \pm d\omega = (c \pm d)\omega, \qquad c\omega \cdot d\omega = \infty\end{gathered} \tag{30}$$

We put

$$\overline{c\omega} = -c\omega, \qquad \overline{\infty} = \infty \tag{30a}$$

where for the extended set of dual numbers (extended by introducing the numbers $c\omega$, ∞) the identity $\bar{\bar{z}} = z$ and all the rules given by Equations 7 remain valid. The expressions 0/0, ∞/∞, and $0 \cdot \infty$ still have no meaning; however, the fraction $(az + b)/(cz + d)$ for $z = \infty$ is naturally thought of as equal to a/c (cf. footnote 4; for $z = k\omega$ this fraction is equal to $(ak + b\varepsilon)/(ck + d\varepsilon)$, since $k\omega = k/\varepsilon$).

A number $c\varepsilon$ of zero modulus can be characterized by the fact that there exists a *nonzero* dual number z $(=d\varepsilon)$, whose product with the number $c\varepsilon$ is zero:

$$c\varepsilon \cdot d\varepsilon = (cd)\varepsilon^2 = 0 \tag{31}$$

Such numbers are therefore called **divisors of zero**.

---

[12] Here we proceed from the fact that, for example, $(c\omega)/(a + b\varepsilon) = (c/\varepsilon)/(a + b\varepsilon)$ is naturally equated to $c/[\varepsilon(a + b\varepsilon)] = c/(a\varepsilon)$, and $(a + b\varepsilon)/c\omega = (a + b\varepsilon)/(c/\varepsilon)$ is regarded as equal to $[\varepsilon(a + b\varepsilon)]/c = (a\varepsilon)/c$.

Dual numbers with nonzero modulus a can also be expressed in a form similar to the trigonometrical form, Equation 13, of a complex number:

$$a + b\varepsilon = a\left(1 + \frac{b}{a}\varepsilon\right) = r(1 + \varepsilon\varphi) \tag{32}$$

Here, as before, $r = a$ is the modulus of the number $z = a + b\varepsilon$, and the ratio $b/a = \varphi$ is called the **argument** of the number and denoted by arg z (r can be any real nonzero number, φ can be any real number). It is obvious that real numbers $a = a + 0 \cdot \varepsilon$ are characterized by the argument being zero; conjugate dual numbers $z = a + b\varepsilon$ and $\bar{z} = a - b\varepsilon$ have the same modulus r and opposite arguments φ and $-\varphi$.

The form for dual numbers, Equation 32, is very convenient when these numbers have to be multiplied or divided. In fact,

$$\begin{aligned} r(1 + \varepsilon\varphi) \cdot r_1(1 + \varepsilon\varphi_1) &= rr_1(1 + \varepsilon\varphi + \varepsilon\varphi_1 + \varepsilon^2\varphi\varphi_1) \\ &= rr_1[1 + \varepsilon(\varphi + \varphi_1)] \end{aligned} \tag{33}$$

hence, *the modulus of the product of two dual numbers is equal to the product of the moduli of the factors,*[13] *and the argument of the product is the sum of the arguments* (cf. Section 1). It follows that *the modulus of the quotient of two dual numbers is equal to the quotient of the moduli of these numbers, and the argument of the quotient is equal to the difference of the corresponding arguments*:

$$\frac{z_1}{z} = \frac{r_1(1 + \varepsilon\varphi_1)}{r(1 + \varepsilon\varphi)} = \frac{r_1}{r}\,[1 + \varepsilon(\varphi_1 - \varphi)] \tag{34}$$

Finally, from these rules we deduce the laws for raising a dual number to any power and taking any root of it:

$$\begin{aligned} [r(1 + \varepsilon\varphi)]^n &= r^n(1 + \varepsilon n\varphi) \\ [r(1 + \varepsilon\varphi)]^{1/n} &= r^{1/n}\left(1 + \varepsilon\,\frac{\varphi}{n}\right) \end{aligned} \tag{35}$$

---

[13] This statement remains valid in the case in which the modulus of one or both factors is equal to zero; for, if $|z| = 0$, then $|zz_1| = 0$, and so, for example, $c\varepsilon(a + b\varepsilon) = (ac)\,\varepsilon$.

(From the latter formula it follows that a root of odd order of a dual number, if $r \neq 0$, is defined uniquely; a root of even order does not exist if $r < 0$, and it has two values if $r > 0$.[14])

**§5. Double Numbers

In complete analogy with everything explained above, we call two double numbers z and $\bar{z}$ **conjugate** if they have the form

$$z = a + be \quad \text{and} \quad \bar{z} = a - be$$

The sum $z + \bar{z} = 2a$ and the product $z \cdot \bar{z} = a^2 - b^2$ of conjugate double numbers are real; the square root of the number $|z\bar{z}| = |a^2 - b^2|$, whose sign agrees with the sign of the larger in absolute value of the real numbers a and b, is called the **modulus** of the number $z = a + be$ and is denoted by $|z|$. It is easy to check that all the Equations 7 remain valid for double numbers; it is also clear that the equation $z = \bar{z}$ characterizes the real numbers $a + 0 \cdot e = a$, and the equation $z = -\bar{z}$ characterizes the purely imaginary numbers $0 + be = be$.

Addition, subtraction, multiplication, and division of double numbers are defined by the formulae

$$(a + be) \pm (c + de) = (a \pm c) + (b \pm d)e$$

$$(a + be) \cdot (c + de) = (ac + bd) + (ad + bc)e$$

$$\frac{c + de}{a + be} = \frac{(c + de)(a - be)}{(a + be)(a - be)} = \frac{(ca - db) + (-cb + da)e}{a^2 - b^2}$$

$$= \frac{ca - db}{a^2 - b^2} + \frac{-cb + da}{a^2 - b^2}e \tag{36}$$

Hence it follows that here also division by $z = a + be$ is possible only in those cases in which $|z| = |a^2 - b^2|^{1/2} \neq 0$. Double

[14] It is not difficult to see that a root of integral order $n > 1$ of a dual number $z = b\varepsilon$, whose modulus $|z|$ is equal to zero (i.e., a root of a divisor of zero), cannot be extracted.

numbers $a \pm ae$, whose moduli are zero, are called **divisors of zero**; we may observe that

$$(a + ae) \cdot (b - be) = ab(1 + e)(1 - e) = 0$$

In some cases it is convenient to regard the quotients $1/(1 + e) = \omega_1$, $1/(1 - e) = \omega_2$, and $1/0 = \infty$ as numbers of a different character; here it is necessary to extend the idea of a double number by introducing, in addition, the products $c\omega_1$ and $c\omega_2$ of the new numbers ω_1 and ω_2 with all possible real numbers c, and the quotients $\omega_1/\omega_2 = (1 - e)/(1 + e) = \sigma_1$ and $\omega_2/\omega_1 = (1 + e)/(1 - e) = \sigma_2$. The rules for division by the symbols $c\omega_1$, $c\omega_2$, ∞, σ_1, and σ_2 are defined by Equations 8 and a series of relations like Equations 30; for example,[15]

$$(a + be) \pm c\omega_1 = (\pm c)\omega_1$$

$$(a + be) \cdot c\sigma_2 = (a + b)c\sigma_2$$

$$\frac{a + be}{c\omega_2} = \frac{a - b}{c}(1 - e), \qquad \frac{a + be}{c\sigma_1} = \frac{a + b}{c}\sigma_2 \tag{37}$$

$$a\omega_1 \cdot b\omega_2 = \infty, \qquad a\omega_1 \cdot b\sigma_2 = ab\omega_2, \qquad a\omega_1 \cdot b\omega_1 = \frac{ab}{2}\omega_1$$

and so on. It is natural to put

$$\overline{c\omega_1} = c\omega_2, \quad \overline{c\omega_2} = c\omega_1, \quad \bar{\sigma}_1 = \sigma_2, \quad \bar{\sigma}_2 = \sigma_1, \quad \overline{\infty} = \infty \tag{37a}$$

---

[15] Since we can naturally say that, for example,

$$(a + be)c\sigma_2 = \frac{(a + be)(c + ce)}{1 - e} = \frac{(a + b)c(1 + e)}{1 - e} = (a + b)c\sigma_2$$

$$\frac{a + be}{c\omega_2} = \frac{(a + be)(1 - e)}{c} = \frac{a - b}{c}(1 - e)$$

$$\frac{a + be}{c\sigma_1} = \frac{(a + be)(1 + e)}{c(1 - e)} = \frac{(a + b)(1 + e)}{c(1 - e)} = \frac{a + b}{c}\sigma_2$$

$$a\omega_1 \cdot b\omega_1 = \frac{ab}{(1 + e)^2} = \frac{ab}{2(1 + e)} = \frac{ab}{2}\omega_1$$

which ensures that the identity $\bar{\bar{z}} = z$ and all the Equations 7 are satisfied for the set of double numbers, extended as shown above.

Double numbers with nonzero modulus may also be expressed in a form analogous to the Equations 13 and 32 for writing ordinary complex numbers and dual numbers. Let $r = \pm \mid a^2 - b^2 \mid^{1/2}$ be the modulus $\mid z \mid$ of a double number; then (cf. Equations 13 and 13a),

$$z = a + be = r\left(\frac{a}{r} + \frac{b}{r}e\right)$$

From the definition of modulus it follows that $(a/r)^2 - (b/r)^2 = \pm 1$ and that the larger (in absolute value) of the fractions a/r and b/r is positive. Hence it follows that

$$\frac{a}{r} = \cosh\varphi, \quad \frac{b}{r} = \sinh\varphi, \qquad \text{or} \qquad \frac{a}{r} = \sinh\varphi, \quad \frac{b}{r} = \cosh\varphi \tag{38}$$

where φ is a certain number (defined by Equations 38) and $\cosh\varphi$ and $\sinh\varphi$ are the *hyperbolic cosine* and the *hyperbolic sine* of the argument φ.[16]

Thus we have

$$z = r(\cosh\varphi + e\sinh\varphi) \qquad \text{or} \qquad z = r(\sinh\varphi + e\cosh\varphi) \tag{39}$$

The quantity φ is called the **argument** of the double number z and is denoted by $\arg z$.

The form given by Equations 39 for writing double numbers is very convenient when it is necessary to multiply two or more double numbers. In fact, from the addition formulae for hyperbolic functions it follows that

$$\begin{aligned} &r(\cosh\varphi + e\sinh\varphi)\cdot r_1(\cosh\varphi_1 + e\sinh\varphi_1) \\ &\qquad = rr_1[\cosh(\varphi + \varphi_1) + e\sinh(\varphi + \varphi_1)] \\ &r(\sinh\varphi + e\cosh\varphi)\cdot r_1(\sinh\varphi_1 + e\cosh\varphi_1) \\ &\qquad = rr_1[\cosh(\varphi + \varphi_1) + e\sinh(\varphi + \varphi_1)] \\ &r(\cosh\varphi + e\sinh\varphi)\cdot r_1(\sinh\varphi_1 + e\cosh\varphi_1) \\ &\qquad = rr_1[\sinh(\varphi + \varphi_1) + e\cosh(\varphi + \varphi_1)] \end{aligned} \tag{40}$$

---

[16] For these functions see, for example, V. G. Shervatov, *Hyperbolic Functions* (D. C. Heath and Company, Boston), 1963.

Thus, *the modulus of the product of two double numbers is equal to the product of the moduli of the factors, and the argument of the product is equal to the sum of the arguments*; here *the product has the form of the first or second of Equations* 39, *depending on whether the factors have the same or different forms*. The rule for dividing double numbers follows immediately from Equations 40:

$$\begin{aligned}\frac{r_1(\cosh\varphi_1 + e\sinh\varphi_1)}{r(\cosh\varphi + e\sinh\varphi)} &= \frac{r_1(\sinh\varphi_1 + e\cosh\varphi_1)}{r(\sinh\varphi + e\cosh\varphi)}\\ &= \frac{r_1}{r}[\cosh(\varphi_1-\varphi) + e\sinh(\varphi_1-\varphi)]\\ \frac{r_1(\sinh\varphi_1 + e\cosh\varphi_1)}{r(\cosh\varphi + e\sinh\varphi)} &= \frac{r_1(\cosh\varphi_1 + e\sinh\varphi_1)}{r(\sinh\varphi + e\cosh\varphi)}\\ &= \frac{r_1}{r}[\sinh(\varphi_1-\varphi) + e\cosh(\varphi_1-\varphi)]\end{aligned} \tag{41}$$

From Equations 40 we obtain also the rules for raising a double number to any positive integral power n and for extracting an nth root of a double number:

$$\begin{aligned}[r(\cosh\varphi + e\sinh\varphi)]^n &= r^n(\cosh n\varphi + e\sinh n\varphi)\\ [r(\sinh\varphi + e\cosh\varphi)]^n &= r^n(\sinh n\varphi + e\cosh n\varphi) \quad \text{if } n \text{ is odd,}\\ &= r^n(\cosh n\varphi + e\sinh n\varphi) \quad \text{if } n \text{ is even;}\\ [r(\cosh\varphi + e\sinh\varphi)]^{1/n} &= r^{1/n}\left(\cosh\frac{\varphi}{n} + e\sinh\frac{\varphi}{n}\right) \text{ if } n \text{ is odd,}\\ &= r^{1/n}\left(\cosh\frac{\varphi}{n} + e\sinh\frac{\varphi}{n}\right)\\ &\quad \text{or}\\ &\quad r^{1/n}\left(\sinh\frac{\varphi}{n} + e\cosh\frac{\varphi}{n}\right) \text{ if } n \text{ is even,}\\ [r(\sinh\varphi + e\cosh\varphi)]^{1/n} &= r^{1/n}\left(\sinh\frac{\varphi}{n} + e\cosh\frac{\varphi}{n}\right) \text{ if } n \text{ is odd;}\\ &\qquad \text{does not exist if } n \text{ is even}\end{aligned} \tag{42}$$

**§6. Hypercomplex Numbers

The most general complex numbers considered above are formed in the following way. To the set of real numbers we adjoin a complex unit E, whose square is by definition equal to

$$E^2 = -pE - q \tag{43}$$

Further, we consider all possible sums of the Form 23, for which addition, subtraction, and multiplication are defined by Equations 24.

As a generalization of complex numbers we have the so-called **hypercomplex numbers**, obtained by adjoining to the set of real numbers certain complex units $E_1, E_2, \ldots, E_n$; these hypercomplex numbers have the form [¶A][16a]

$$Z = a_0 + a_1E_1 + a_2E_2 + \cdots + a_nE_n, \qquad a_0, a_1, \ldots, a_n \text{ real} \tag{44}$$

The sum, difference, and product of hypercomplex numbers are defined by the formulae

$$\begin{aligned}
&(a_0 + a_1E_1 + \cdots + a_nE_n) \pm (b_0 + b_1E_1 + \cdots + b_nE_n)\\
&\qquad = (a_0 \pm b_0) + (a_1 \pm b_1)E_1 + \cdots + (a_n \pm b_n)E_n\\
&(a_0 + a_1E_1 + \cdots + a_nE_n) \cdot (b_0 + b_1E_1 + \cdots + b_nE_n)\\
&\qquad = a_0b_0 + a_0b_1E_1 + \cdots + a_0b_nE_n\\
&\qquad\quad + a_1b_0E_1 + a_1b_1E_1^2 + \cdots + a_1b_nE_1E_n\\
&\qquad\quad + \cdots + a_nb_0E_n + a_nb_1E_nE_1 + \cdots + a_nb_nE_n^2
\end{aligned} \tag{45}$$

For the product of two hypercomplex numbers to be a number of the same kind, it is necessary to give a *multiplication table of the complex units* $E_1, E_2, \ldots, E_n$[17]:

$$\begin{gathered}
E_i \cdot E_j = p_0^{(i,j)} + p_1^{(i,j)}E_1 + \cdots + p_n^{(i,j)}E_n,\\
i, j = 1, 2, \ldots, n
\end{gathered} \tag{46}$$

[16a] These bracketed letters refer to the Addenda at the end of the book.

[17] Sometimes hypercomplex numbers are defined in such a way that the number 1 does not appear among them; in such case a hypercomplex number is defined as

$$Z = a_0E_0 + a_1E_1 + \cdots + a_nE_n \tag{44'}$$

That is, we must give $n^2(n+1)$ real numbers $p_k^{(i,j)}$, where $i, j = 1, 2,\ldots, n$, and $k = 0, 1,\ldots, n$. If $n = 1$, then $n^2(n+1) = 2$; in this case the role of the numbers $p_0^{(1,1)}$ and $p_1^{(1,1)}$ is played by the numbers $-q$ and $-p$, where p and q are the coefficients of Equation 1, which is satisfied by a (unique) complex unit E.

When $n > 1$ the system of hypercomplex numbers is generally *noncommutative* and *nonassociative*; that is, as a rule the product $Z_1 \cdot Z_2$ is different from the product $Z_2 \cdot Z_1$, and the product $(Z_1 \cdot Z_2)Z_3$ is different from the product $Z_1(Z_2 \cdot Z_3)$. The requirements of commutativity and associativity impose on the numbers $p_k^{(i,j)}$ certain conditions (thus, for example, if $E_iE_j = E_jE_i$, then obviously $p_k^{(i,j)} = p_k^{(j,i)}$ for any k), which in general cannot be satisfied.

The general theory of hypercomplex numbers constitutes a branch of algebra. As far as geometry is concerned, the systems of hypercomplex numbers with the most important applications [¶B] are those in which every number Z has a unique conjugate $\bar{Z}$ and a modulus $|Z|$, which have the same properties as those of the ordinary complex and double numbers (that is, $\bar{\bar{Z}} = Z$, $|Z|^2 = Z\bar{Z}$ is real and satisfies the additional condition of "nondegeneration," and $|Z_1Z_2| = |Z_1| \cdot |Z_2|$, and identities of the forms given in Equations 7 are satisfied), and the limiting cases of degenerate complex numbers with the same properties. It can be shown that the *number n of complex units* for such systems of hypercomplex

---

and the multiplication table of the complex units is defined by the formula

$$E_i \cdot E_j = p_0^{(i,j)}E_0 + p_1^{(i,j)}E_1 + \cdots + p_n^{(i,j)}E_n, \quad i, j = 0, 1,\ldots, n \tag{46'}$$

If the complex unit E_0 is such that $E_0E_i = E_iE_0 = E_i$, where $i = 0, 1,\ldots, n$, then it is possible to denote it by 1 and to write the hypercomplex number in the form given in Equation 44.

The form of writing hypercomplex numbers given in Equation 44′ is convenient also in the case of *double numbers* $z = a + be$, where the two complex units $e_1 = (1 + e)/2$ and $e_2 = (1 - e)/2$ are often taken as a basis. It is clear that any double number can be expressed in the form $z = a_1e_1 + a_2e_2$, where the "multiplication table" of the basic units has the very simple form $e_1^2 = e_1$, $e_2^2 = e_2$, $e_1e_2 = 0$. However, we shall not use this form of writing double numbers in what follows.

numbers *can be equal only to* 1, 3, *or* 7. If $n = 3$, the system must be noncommutative,[17a] and if $n = 7$ it is also nonassociative. If $n = 1$, we obtain the cases, considered above, of ordinary, dual, and double numbers. If $n = 3$, our requirements lead to the well-known system of **quaternions**, introduced by the famous Irish mathematician of the nineteenth century W. R. Hamilton (1805–1865),

$$\begin{gathered} Z = a + bi + cj + dk, \qquad i^2 = j^2 = k^2 = -1 \\ ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j \\ \bar{Z} = a - bi - cj - dk, \qquad |Z|^2 = Z\bar{Z} = a^2 + b^2 + c^2 + d^2 \end{gathered} \tag{47}$$

and also the systems related to quaternions, the *pseudoquaternions*, *degenerate quaternions*, *degenerate pseudoquaternions*, and *doubly degenerate quaternions*,

$$\begin{aligned} &Z = a + bi + ce + df, \qquad &&Z = a + bi + c\varepsilon + d\eta \\ &Z = a + be + c\varepsilon + d\zeta, \qquad &&Z = a + b\varepsilon + c\eta + d\zeta \end{aligned} \tag{48}$$

of which the multiplication tables for the complex units have the form

$$\begin{array}{c|ccc} & i & e & f \\ \hline i & -1 & f & -e \\ e & -f & 1 & -i \\ f & e & i & 1 \end{array} \qquad \begin{array}{c|ccc} & i & \varepsilon & \eta \\ \hline i & -1 & \eta & -\varepsilon \\ \varepsilon & -\eta & 0 & 0 \\ \eta & \varepsilon & 0 & 0 \end{array}$$

$$\begin{array}{c|ccc} & e & \varepsilon & \zeta \\ \hline e & 1 & \zeta & \varepsilon \\ \varepsilon & -\zeta & 0 & 0 \\ \zeta & -\varepsilon & 0 & 0 \end{array} \qquad \begin{array}{c|ccc} & \varepsilon & \eta & \zeta \\ \hline \varepsilon & 0 & \zeta & 0 \\ \eta & -\zeta & 0 & 0 \\ \zeta & 0 & 0 & 0 \end{array} \tag{49}$$

(that is, $i^2 = ii = -1$, $ie = f$, $if = -e$, and so on), and in which the conjugate hypercomplex number and modulus of a number

[17a] For noncommutative systems of hypercomplex numbers, the rule $\overline{Z_1Z_2} = \bar{Z}_1 \cdot \bar{Z}_2$ (see formulae 7 above) is replaced by $\overline{Z_1Z_2} = \bar{Z}_2 \cdot \bar{Z}_1$ (and correspondingly $\overline{Z_1/Z_2} = \bar{Z}_1/\bar{Z}_2$ by $\overline{Z_1Z_2^{-1}} = \bar{Z}_2^{-1}\bar{Z}_1$, where $ZZ^{-1} = Z^{-1}Z = 1$).

are defined by the formulae

$$
\begin{array}{ll}
\bar{Z} = a - bi - ce - df, & |Z|^2 = Z\bar{Z} = a^2 + b^2 - c^2 - d^2 \\
\bar{Z} = a - bi - c\varepsilon - d\eta, & |Z|^2 = Z\bar{Z} = a^2 + b^2 \\
\bar{Z} = a - be - c\varepsilon - d\zeta, & |Z|^2 = Z\bar{Z} = a^2 - b^2 \\
\bar{Z} = a - b\varepsilon - c\eta - d\zeta, & |Z|^2 = Z\bar{Z} = a^2
\end{array}
\tag{50}
$$

Finally, if $n = 7$ we obtain the system of **octaves**, first considered by the well-known English geometer of the end of the nineteenth century A. Cayley (1821–1895),

$$Z = a_0 + a_1 i_1 + a_2 i_2 + a_3 i_3 + a_4 i_4 + a_5 i_5 + a_6 i_6 + a_7 i_7 \tag{51}$$

of which the multiplication table for the complex units is

	i_1	i_2	i_3	i_4	i_5	i_6	i_7
i_1	-1	i_3	$-i_2$	i_5	$-i_4$	$-i_7$	i_6
i_2	$-i_3$	-1	i_1	i_6	i_7	$-i_4$	$-i_5$
i_3	i_2	$-i_1$	-1	i_7	$-i_6$	i_5	$-i_4$
i_4	$-i_5$	$-i_6$	$-i_7$	-1	i_1	i_2	i_3
i_5	i_4	$-i_7$	i_6	$-i_1$	-1	$-i_3$	i_2
i_6	i_7	i_4	$-i_5$	$-i_2$	i_3	-1	$-i_1$
i_7	$-i_6$	i_5	i_4	$-i_3$	$-i_2$	i_1	-1

(52)

and of which the definitions of conjugate number and modulus are

$$
\begin{array}{l}
\bar{Z} = a_0 - a_1 i_1 - a_2 i_2 - a_3 i_3 - a_4 i_4 - a_5 i_5 - a_6 i_6 - a_7 i_7 \\
|Z|^2 = Z\bar{Z} = a_0^2 + a_1^2 + a_2^2 + a_3^2 + a_4^2 + a_5^2 + a_6^2 + a_7^2
\end{array}
\tag{53}
$$

and also some systems of hypercomplex numbers related to octaves, which can be called *pseudooctaves*, *degenerate octaves*, *degenerate pseudooctaves*, *doubly degenerate octaves*, *doubly degenerate pseudooctaves*, and *triply degenerate octaves* [¶c]. All these systems of hypercomplex numbers have applications to geometry; however, an account of these questions would take us too far.[18]

[18] See, for example, the article refered to in footnote 11 on p. 14, and also an article (in Russian) by B. A. Rozenfel'd and I. M. Yaglom, "On the geometries of the simplest algebras," *Matematicheskii sbornik*, Vol. 28, No. 1 (1951), pp. 205–216, and the book (in Russian) by B. A. Rozenfel'd, *Non-Euclidean Geometries*, Chap. VI (Gostekhizdat, Moscow), 1955, where the algebraic literature referred to here is also mentioned.

CHAPTER II

Geometrical Interpretation of Complex Numbers

§7. Ordinary Complex Numbers as Points of a Plane[18a]

The development of the theory of complex numbers is very closely connected with the geometrical interpretation of ordinary complex numbers as points of a plane, which apparently was first mentioned by the Danish surveyor K. Wessel (1745–1818) but became widely known chiefly through the works of the famous mathematicians K. F. Gauss (1777–1855) and A. Cauchy (1789–1857). This interpretation arises from the fact that the point of a plane with rectangular cartesian coordinates x and y or polar coordinates r and φ corresponds to the *complex number* (see Figure 1):

$$z = x + iy = r(\cos\varphi + i\sin\varphi)$$

Here, obviously, real numbers $z = x + 0 \cdot i = r(\cos 0 + i \sin 0)$ correspond to points of the x axis, the **real axis** o; numbers of modulus $r = 1$ correspond to points of the circle S with center at O and radius 1, the **unit circle**. Opposite complex numbers $z = x + iy$ and $-z = -x - iy$ correspond to points symmetrical about the point O (the number 0 corresponds to the origin

[18a] The contents of Sections 7, 8, 13 and 14 have many points of contact with the books of R. Deaux, *Introduction to the Geometry of Complex Numbers* (F. Ungar Publishing Co.), 1956 and H. Schwerdtfeger, *Geometry of Complex Numbers* (University of Toronto, Oliver and Boyd), 1962; the latter book also contains much material which could supplement the contents of Section 11 of the present book.

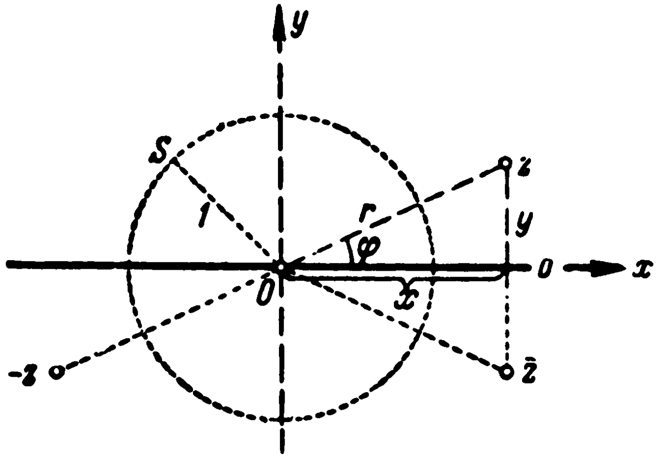


FIG. 1

O); conjugate complex numbers $z = x + iy = r(\cos\varphi + i\sin\varphi)$ and

$$z = x - iy = r[\cos(-\varphi) + i\sin(-\varphi)]$$

correspond to points symmetrical about the line o (throughout this book we shall consider a line to mean a *straight line*). Henceforth we shall often denote the point that corresponds to a complex number z by the same letter z; in this case, the equations

$$z' = -z \tag{1}$$

$$z' = \bar{z} \tag{2}$$

can be regarded as definite **point transformations** which associate each point z with a new point z'. Equations 1 and 2 represent the **symmetry** (*half-turn*) **about the point** O and the **symmetry** (*reflection*) **about the line** o.

Let $q = a + ib$ and $p = t(\cos\alpha + i\sin\alpha)$ be fixed complex numbers (points). The equation

$$z' = z + q, \quad \text{i.e.,} \quad x' + iy' = (x + iy) + (a + ib) \tag{3}$$

means that $x' = x + a$ and $y' = y + b$, that is, that the vector $\overrightarrow{zz'}$ (the vector beginning at the point z and ending at the point z') is equal to the vector $\overrightarrow{Oq}$ (*the geometrical meaning of addition of*

complex numbers). Therefore Equation 3 defines a **translation** (*parallel displacement*) of the plane by the vector $\overrightarrow{Oq}$; see Figure 2.

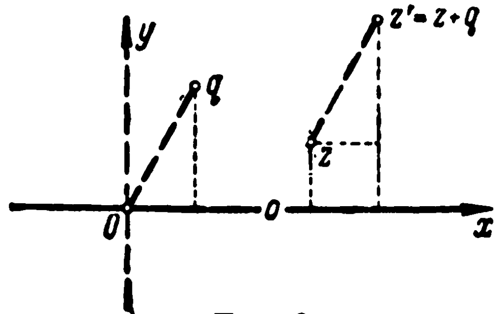


FIG. 2

The equation $z' = pz$, i.e.,

$$r'(\cos\varphi' + i\sin\varphi') = t(\cos\alpha + i\sin\alpha) \times r(\cos\varphi + i\sin\varphi) \qquad (4)$$

means that $r' = tr$ and $\varphi' = \varphi + \alpha$ (see Figure 3); that is,

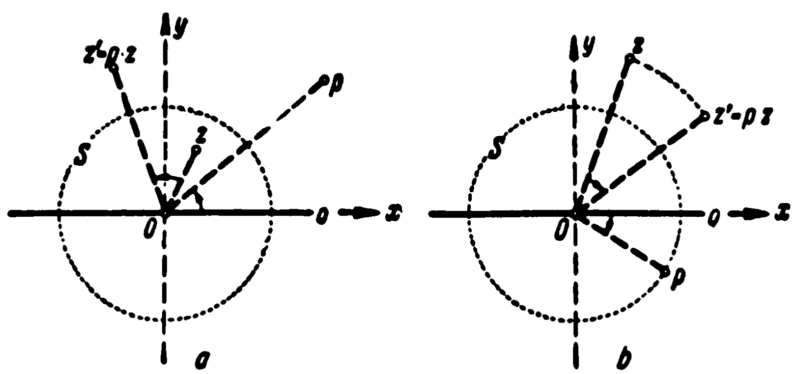


FIG. 3

$(O, z') = t \cdot (O, z)$, where (O, z) is the distance between the points O and z, and $\angle\{Oz, Oz'\} = \alpha$ (*the geometrical meaning of multiplication of complex numbers*). The symbol $\angle\{Oz, Oz'\}$ denotes the so-called **oriented angle** (the braces indicate orientation) between the rays Oz and Oz', the angle through which Oz must be rotated *counterclockwise* to yield Oz' (note that if the rotation is clockwise, a minus sign is attached to the angle; the oriented angle between two rays is defined apart from the addition of a multiple of 2π). Thus, Equation 4 defines a so-called **spiral similarity** of the plane, which consists of a **rotation** about O through an angle α (in the counterclokwise sense) and

a **dilatation** (*homothety*) with center O and ratio t. In particular, if the modulus t of the complex number p is equal to 1, the transformation given by Equation 4 is a *rotation* through an angle α (Figure 3).

Every motion of the plane can be represented as a rotation about the fixed point O, together with a translation, or as the symmetry about the fixed line o, together with a rotation about the chosen point O and a translation.[19] It follows that *every motion of the plane can be written in one of the forms*[20]

$$z' = pz + q, \qquad |p| = 1 \tag{5}$$

$$z' = p\bar{z} + q, \qquad |p| = 1 \tag{5a}$$

It is obvious that the **distance** $d = (z_1, z_2)$ *between two points* z_1 *and* z_2 *of the plane is equal to the modulus* $|w| = |z_2 - z_1|$ *of the complex number* $w = z_2 - z_1$ (since the vector $\overrightarrow{z_1z_2}$ is equal to the vector $\overrightarrow{Ow}$); see Figure 4. Put differently,

$$d = |z_2 - z_1|, \qquad d^2 = (z_2 - z_1)(\bar{z}_2 - \bar{z}_1) \tag{6}$$

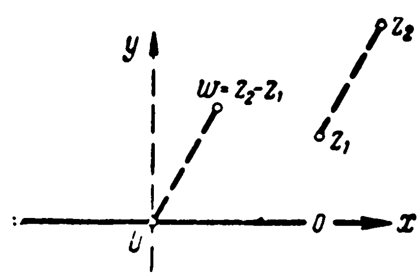


FIG. 4

Further, it is obvious that the **angle** δ_0 *between two lines intersecting at the origin* O *and defined by the points* z_1^0 *and* z_2^0 (Figure 5) *is equal to*

$$\delta_0 = \arg \frac{z_2^0}{z_1^0} \qquad (= \arg z_2^0 - \arg z_1^0) \tag{7}$$

[19] See, for example, I. M. Yaglom, *Geometric Transformations* (Random House), 1962.

[20] Similarly, it can be shown that *every similarity of the plane can be written in the form* $z' = pz + q$ or $z' = p\bar{z} + q$.

This allows us to define the angle δ between lines intersecting at an arbitrary point z_0 and passing through the points z_1 and z_2. The simplest motion which takes z_0 into O is the translation

$$z' = z - z_0$$

This translation takes the points z_1 and z_2 into the points $z_1^0 = z_1 - z_0$ and $z_2^0 = z_2 - z_0$; see Figure 5. Hence it follows that the angle δ is equal to the angle δ_0 between the lines which intersect at O and pass through the points z_1^0 and z_2^0; that is

$$\delta = \arg \frac{z_2 - z_0}{z_1 - z_0} \qquad (= \arg(z_2 - z_0) - \arg(z_1 - z_0)) \quad (8)$$

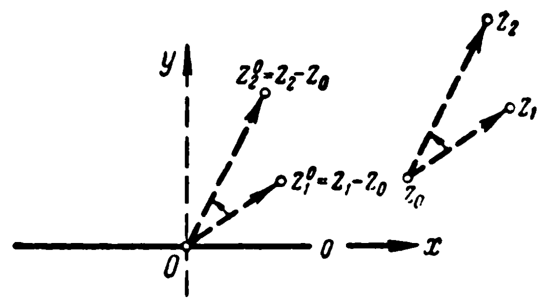


FIG. 5

We shall call the complex number

$$V(z_2, z_1, z_0) = \frac{z_2 - z_0}{z_1 - z_0}$$

the **ratio of the three points** (three complex numbers) z_2, z_1, z_0. Thus, *the angle δ between the lines which intersect at a point z_0 and pass through the points z_1 and z_2 is equal to the argument of the ratio $V(z_2, z_1, z_0)$ of the points z_2, z_1, z_0.*

We note that δ is the angle through which the *ray* $\overrightarrow{z_0 z_1}$ of the first line must be rotated counterclockwise to be brought into line with the *ray* $\overrightarrow{z_0 z_2}$ of the second line. A line of whose two directions one is selected is called an **oriented line** or **axis** (this direction can be given by indicating the ray defined on the line, and on diagrams it is usually denoted by an arrow); the selected direction of an oriented line is often called **positive.** The oriented line passing through the two points z_1 and z_2, the positive direction of which coincides with the direction from

z_1 to z_2, is denoted by $[z_1 z_2]$. The **oriented angle** $\angle\{l_1, l_2\}$ between the oriented lines l_1 and l_2 is defined as the angle through which the line l_1 must be rotated counterclockwise for its positive direction to coincide with the positive direction of the line l_2; this angle is given, apart from the addition of a multiple of 2π. Thus, δ *is the oriented angle between the oriented lines* $[z_0 z_1]$ *and* $[z_0 z_2]$, or $\delta = \angle\{[z_0 z_1], [z_0 z_2]\}$.

The condition for three points z_0, z_1, *and* z_2 *to lie on one line is that the angle* $\angle\{[z_1 z_2], [z_0 z_2]\}$ *is equal to* 0 *or* π or, by virtue of Equation 8, that *the ratio* $V(z_0, z_1, z_2) = (z_0 - z_2)/(z_1 - z_2)$ *of these three points is real.* This condition may also be written

$$\frac{z_0 - z_2}{z_1 - z_2} = \frac{\bar{z}_0 - \bar{z}_2}{\bar{z}_1 - \bar{z}_2} \tag{9}$$

Hence it follows that the line l through the points z_1 and z_2 is the **locus** of those points z for which

$$\frac{z - z_2}{z_1 - z_2} = \frac{\bar{z} - \bar{z}_2}{\bar{z}_1 - \bar{z}_2} \tag{10}$$

In other words, we can say that the *equation of this line* has the form of Equation 10 or of

$$(\bar{z}_1 - \bar{z}_2)z - (z_1 - z_2)\bar{z} + (z_1\bar{z}_2 - \bar{z}_1 z_2) = 0 \tag{10a}$$

Thus, *the equation of any line can be written in the form*[21]

$$Bz - \bar{B}\bar{z} + C = 0, \qquad C \text{ purely imaginary} \tag{11}$$

---

[21] It is obvious that the angle $\angle\{l, o\}$, which the line l described by Equation 11 makes with the real axis o, is equal to $\arg B = \arg(\bar{z}_1 - \bar{z}_2) = -\arg(z_1 - z_2)$; further, since $|C| = |Bz - \bar{B}\bar{z}| \leqslant 2|B|\cdot|z|$, where $|Bz - \bar{B}\bar{z}| = 2|B|\cdot|z|$ only if Bz is purely imaginary (that is, if $\arg B + \arg z = \pm\arg i$), we have, for points z of the line l, $|z| \leqslant |C|/2|B|$, whence it follows that $|C|/2|B|$ is the distance of l from the origin.

Equation 11 can also be deduced from the fact that by the motion given by Equation 5 the line l can be taken into the real axis $z - \bar{z} = 0$. Hence we obtain $(pz + q) - (\bar{p}\bar{z} + \bar{q}) = 0$ or $Bz - \bar{B}\bar{z} + C = 0$, where $B = p$ and $C = q - \bar{q}$. This deduction also enables us to conclude that $\angle\{l, o\} = \arg p = \arg B$ and that $|C|/2|B|$ is the distance of l from the origin.

Equation 11 is often written in the form $bz + \bar{b}\bar{z} + c = 0$, where c is real; here $b = Bi$ and $c = Ci$.

It is not difficult to show that, conversely, *every equation of this form gives some line* (passing through the points z_1 and z_2, such that $z_1 - z_2 = \bar{B}$ and $z_1\bar{z}_2 - \bar{z}_1 z_2 = C$).

The condition for four points z_0, z_1, z_2, and z_3 to lie on one circle (or line) is that the difference of the angles,

$$\angle\{[z_0 z_2], [z_1 z_2]\} - \angle\{[z_0 z_3], [z_1 z_3]\},$$

is equal to 0 or π (see Figure 6) or, by virtue of Equation 8, that the number

$$\frac{V(z_0, z_1, z_2)}{V(z_0, z_1, z_3)} = \frac{z_0 - z_2}{z_1 - z_2} : \frac{z_0 - z_3}{z_1 - z_3}$$

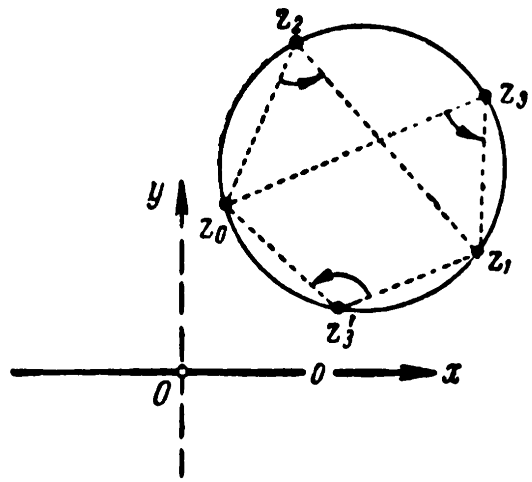


FIG. 6

is real. This ratio of the ratios of the two triads of points (z_0, z_1, z_2 and z_0, z_1, z_3) is called the **cross-ratio of the four points** z_0, z_1, z_2, z_3, and is denoted by $W(z_0, z_1, z_2, z_3)$. Thus, *the condition for four points z_0, z_1, z_2, and z_3 to lie on one circle (or line) is that their cross-ratio,*

$$W(z_0, z_1, z_2, z_3) = \frac{z_0 - z_2}{z_1 - z_2} : \frac{z_0 - z_3}{z_1 - z_3}$$

is real. This condition may be written:

$$\frac{z_0 - z_2}{z_1 - z_2} : \frac{z_0 - z_3}{z_1 - z_3} = \frac{\bar{z}_0 - \bar{z}_2}{\bar{z}_1 - \bar{z}_2} : \frac{\bar{z}_0 - \bar{z}_3}{\bar{z}_1 - \bar{z}_3} \tag{12}$$

From this it follows that the equation of the circle (or line) S passing through the points z_1, z_2, and z_3 has the form

$$\frac{z - z_2}{z_1 - z_2} : \frac{z - z_3}{z_1 - z_3} = \frac{\bar{z} - \bar{z}_2}{\bar{z}_1 - \bar{z}_2} : \frac{\bar{z} - \bar{z}_3}{\bar{z}_1 - \bar{z}_3}, \tag{13}$$

$$\begin{aligned}(z - z_2)(\bar{z} - \bar{z}_3)[(z_1 - z_3)(\bar{z}_1 - \bar{z}_2)]\\ = (z - z_3)(\bar{z} - \bar{z}_2)[(z_1 - z_2)(\bar{z}_1 - \bar{z}_3)]\end{aligned}$$

That is,

$$Az\bar{z} + Bz - \bar{B}\bar{z} + C = 0$$

where

$$\begin{aligned}A &= (z_1 - z_3)(\bar{z}_1 - \bar{z}_2) - (z_1 - z_2)(\bar{z}_1 - \bar{z}_3)\\ B &= -\bar{z}_3(z_1 - z_3)(\bar{z}_1 - \bar{z}_2) + \bar{z}_2(z_1 - z_2)(\bar{z}_1 - \bar{z}_3)\\ C &= z_2\bar{z}_3(z_1 - z_3)(\bar{z}_1 - \bar{z}_2) - z_3\bar{z}_2(z_1 - z_2)(\bar{z}_1 - \bar{z}_3)\end{aligned}$$

Thus, *the equation of every circle (or line) may be written in the following form*:

$$Az\bar{z} + Bz - \bar{B}\bar{z} + C = 0, \qquad A,\ C \text{ purely imaginary} \tag{14}$$

Conversely, *the locus of points z which satisfy any equation of this form is a circle or a line* (if such points exist).[22]

We know already that Equation 14 *denotes a line if (and only if)*

$$A = 0 \tag{15}$$

---

[22] Equation 14 can also be deduced from the fact that by the translation given by Equation 3 any circle S can be taken into the circle $z\bar{z} = r^2$ with center at origin O (r is the radius of the circle). Hence we obtain the equation of S, $(z + q)(\bar{z} + \bar{q}) - r^2 = 0$ or $az\bar{z} + bz + \bar{b}\bar{z} + c = 0$, where $a = 1$ and $c = q\bar{q} - r^2$ are real, and where $b = \bar{q}$. From this equation it is obvious that the square of the radius r of the circle S is equal to $(b\bar{b} - ac)/a^2$ and its center is the point $-b/a$. (Thus, if $b\bar{b} - ac = 0$, the equation of S is satisfied by the single point $z = -b/a$, and if $b\bar{b} - ac < 0$ it is not satisfied by any point of the plane.)

The equation of S in the form mentioned here is obviously equivalent to Equation 14 (to go over from one of these equations to the other it is sufficient to put $a = Ai$, $b = Bi$, and $c = Ci$).

We note, too, that by virtue of the rules for operating with the symbol ∞, explained above, we must write

$$W(z_1, z_2, z_3, \infty) = \frac{z_1 - z_3}{z_2 - z_3} : \frac{z_1 - \infty}{z_2 - \infty}$$

$$= \frac{z_1 - z_3}{z_2 - z_3} = V(z_1, z_2, z_3)$$

since $(z_1 - \infty)/(z_2 - \infty) = 1$ (cf. footnote 4). Consequently, the cross-ratio $W(z_1, z_2, z_3, \infty)$ is real, if and only if the ordinary ratio $V(z_1, z_2, z_3)$ is real, that is, when the points z_1, z_2, and z_3 lie on one line. Therefore it is appropriate to regard *the* **point at infinity** *of the plane, corresponding to the "number"* ∞ *as lying on all lines* (in certain cases this fictitious point is useful). In fact, if three points z_1, z_2, and z_3 lie on one line, then the line through these points may be defined as the locus of points z such that $W(z_1, z_2, z_3, z)$ is real; however, the "point" ∞ satisfies the latter condition.

*§8. Applications and Examples

The work done so far enables us to use complex numbers to prove numerous theorems dealing with lines and circles. We mention here some examples.

We begin with the following theorem. *Let four circles S_1, S_2, S_3, and S_4 be given in a plane, and let z_1 and w_1 be the points of intersection of S_1 and S_2, z_2 and w_2 those of S_2 and S_3, z_3 and w_3 those of S_3 and S_4, and z_4 and w_4 those of S_4 and S_1. If the points z_1, z_2, z_3, and z_4 lie on one circle (or line) $\sum$, then the points w_1, w_2, w_3, and w_4 lie on one circle (or line) $\sum'$*; see Figure 7.

We use the fact that the points z_1, z_2, w_1, and w_2 lie on one circle S_2, the points z_2, z_3, w_2, and w_3 on the circle S_3, the points z_3, z_4, w_3, and w_4 on the circle S_4, and the points z_4, z_1, w_4,

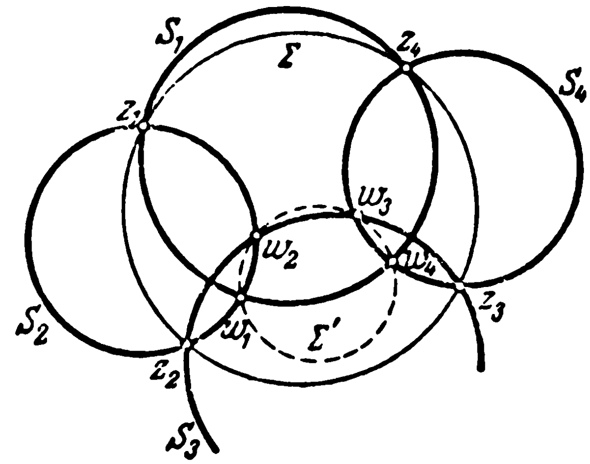


FIG. 7

and w_1 on the circle S_1. It follows that the following four cross-ratios are real:

$$W(z_1, w_2, z_2, w_1) = \frac{z_1 - z_2}{w_2 - z_2} : \frac{z_1 - w_1}{w_2 - w_1}$$

$$W(z_2, w_3, z_3, w_2) = \frac{z_2 - z_3}{w_3 - z_3} : \frac{z_2 - w_2}{w_3 - w_2}$$

$$W(z_3, w_4, z_4, w_3) = \frac{z_3 - z_4}{w_4 - z_4} : \frac{z_3 - w_3}{w_4 - w_3}$$

$$W(z_4, w_1, z_1, w_4) = \frac{z_4 - z_1}{w_1 - z_1} : \frac{z_4 - w_4}{w_1 - w_4}$$

Consequently, the expression

$$\frac{W(z_1, w_2, z_2, w_1) \cdot W(z_3, w_4, z_4, w_3)}{W(z_2, w_3, z_3, w_2) \cdot W(z_4, w_1, z_1, w_4)}$$

$$= \left(\frac{z_1 - z_2}{z_3 - z_2} : \frac{z_1 - z_4}{z_3 - z_4}\right)\left(\frac{w_1 - w_2}{w_3 - w_2} : \frac{w_1 - w_4}{w_3 - w_4}\right)$$

$$= W(z_1, z_3, z_2, z_4) W(w_1, w_3, w_2, w_4)$$

is real. Therefore, since the cross-ratio $W(z_1, z_3, z_2, z_4)$ is real, it follows that the cross-ratio $W(w_1, w_3, w_2, w_4)$ is real, which proves the theorem.

This proposition seems rather elegant, but not particularly promising—an ordinary theorem, of which there are many in elementary geometry. However, the consequences which follow from this simple theorem can surely be called remarkable. As the first of these consequences we mention a whole series of theorems due to the English geometer

W. Clifford. We agree to call *n* lines of a plane *lines in general position* if no two of them are parallel and no three meet in a point. We call the point of intersection of two lines in general position (i.e., intersecting lines) their **central point** (Figure 8a). From three lines in general position it is possible to choose a pair of lines in three different ways; these three pair of lines determine three central points; by taking the circle through them (i.e. the circumcircle of the triangle formed by the three lines) we obtain the **central circle of the three lines** (Figure 8b). Similarly, from four lines in general position it is possible

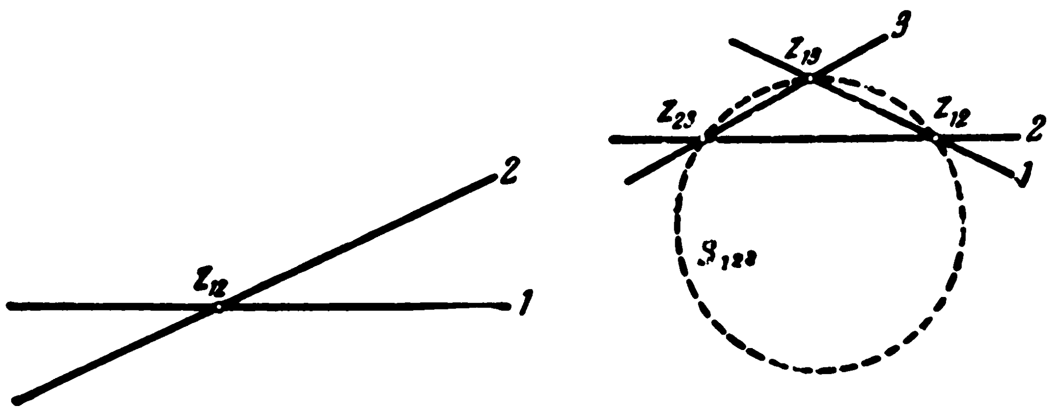


FIG. 8 a. & b.

to choose a triad of lines in four different ways; these four triads determine four central circles, which *always meet in one point*; this point is naturally called the **central point of the four lines** (Figure 8c). From five lines in general position it is possible to choose a tetrad of lines in five different ways; the five tetrads thus obtained determine

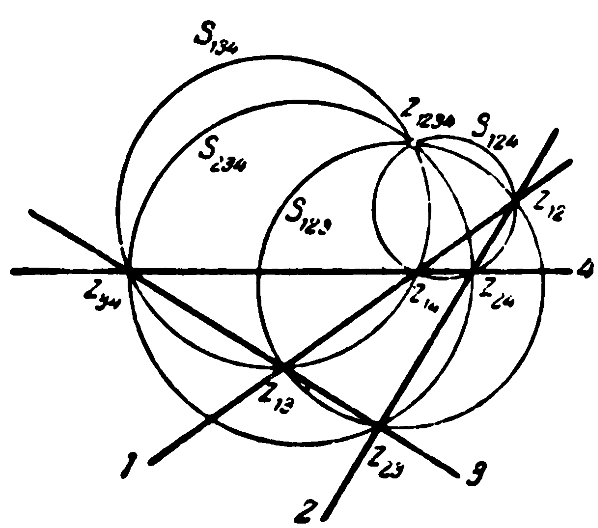


FIG. 8c.

five central points, which *always lie on one circle*, the **central circle of the five lines** (Figure 8d).

We may note that this proposition is equivalent to the following. *If we produce the sides of an arbitrary* (possibly nonconvex or even

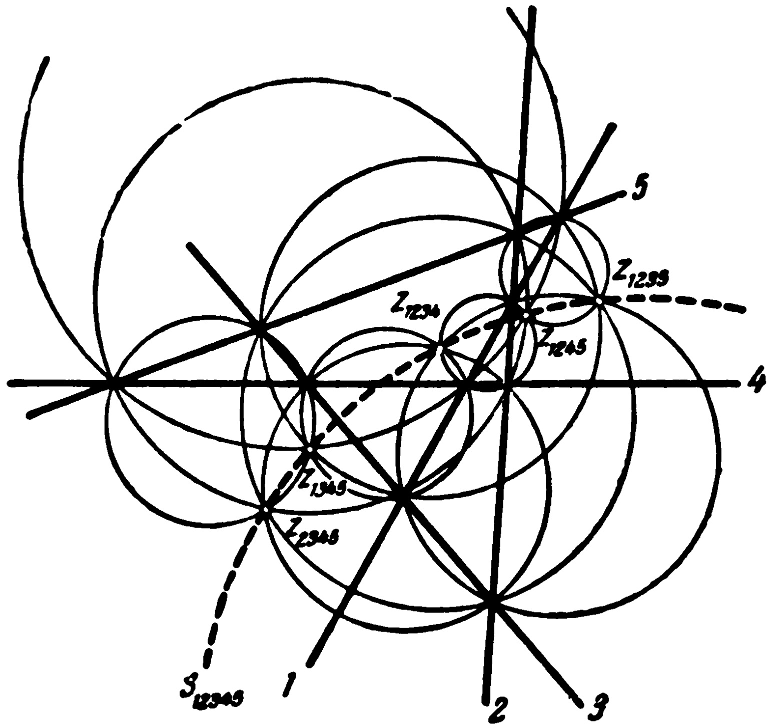


FIG. 8d.

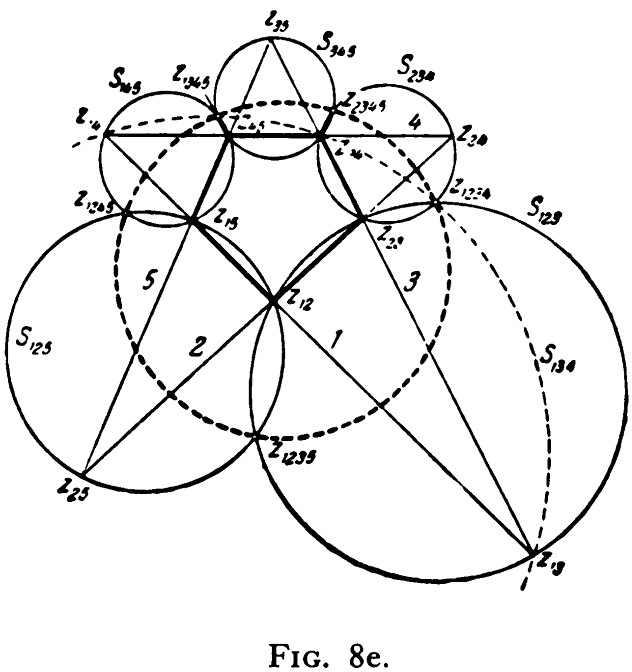


FIG. 8e.

self-intersecting) *pentagon, until each side meets the next but one, and describe the circumcircles of the five triangles thus produced, then the five points of intersection of adjacent circles lie on one circle*, the central circle of the sides of the pentagon (Figure 8e).

In a similar way we can continue this definition as far as we like: *every even number n of lines in general position will determine a central point, the point of intersection of the n central circles of every system of n — 1 of these lines, and every odd number n of lines in general position will determine a central circle, which passes through the n central points of every system of n — 1 of these lines.*

For a proof of this proposition we consider first of all the case of *four* lines 1, 2, 3, and 4. The points of intersection of these lines (the central points of pairs of the lines) we denote by z_{12}, z_{13}, z_{14}, z_{23}, z_{24}, and z_{34}, as shown in Figure 8c; the point of intersection, other than z_{34}, of the circles through the points z_{23}, z_{34}, z_{24}, and z_{13}, z_{34}, z_{14} (the central circles S_{234} and S_{134} of the triads of lines 2, 3, 4 and 1, 3, 4) we denote by z_{1234}. In this case we arrive at a familiar situation. Here the role of the four circles S_1, S_2, S_3, and S_4 is played by the circle S_{234}, the line 2, the line 1, and the circle S_{134}; the role of the points z_1, z_2, z_3, and z_4 is played by the point z_{23}, the point at infinity ∞, through which all lines "pass" (see the end of the preceding section), and the points z_{13} and z_{34}; the role of the points w_1, w_2, w_3, and w_4 is played by the points z_{24}, z_{12}, z_{14}, and z_{1234}. From the fact that the points z_{23}, ∞, z_{13}, and z_{34} lie on one circle or line (in fact, line 3), it follows that the points z_{24}, z_{12}, z_{14}, and z_{1234} also lie on one circle; that is, the central circle S_{124} of the lines 1, 2, and 4, which passes through z_{12}, z_{24}, and z_{14}, also passes through the point z_{1234}. Similarly it may be shown that the central circle S_{123} of the lines 1, 2, and 3 passes through the same point z_{1234} (for a proof of this it is sufficient to take for z_1 and z_3 the points z_{24} and z_{14} and for w_1 and w_3 the points z_{23} and z_{13}), whence it follows that z_{1234} is the central point of the four lines 1, 2, 3, and 4.

We now proceed to the case of *five* lines 1, 2, 3, 4, and 5. We denote the central point of the four lines 1, 2, 3, and 4 by z_{1234}, and similarly for the other tetrads of lines; we denote the point of intersection of the lines 1 and 2 (the central point of these two lines) by z_{12}, and similarly for the other pairs of lines; we denote the central circle of the lines 1, 2, and 3 by S_{123}, and similarly for the other triads of lines (Figure 8e). We must show that the points z_{1234}, z_{1235}, z_{1245}, z_{1345}, and z_{2345} lie on one circle, but for this it is sufficient to ensure that one circle passes through *any four* of these points, for example z_{1234}, z_{1235}, z_{1245}, and z_{1345}. This follows immediately from the theorem proved above, since these four points can be regarded as the points of intersection of the circles S_{134} and S_{123}, S_{123} and S_{125}, S_{125} and S_{145}, and

S_{145} and S_{134}; the second points of intersection of these pairs of circles, points z_{13}, z_{12}, z_{15}, and z_{14}, lie on one "circle or line" (in fact, line 1).

In exactly the same way we may prove the requisite theorems in the case of an arbitrary number n of lines, which are naturally denoted by the numbers 1, 2, 3,..., n. We suppose that for any number m of lines, less than n, our theorem has already been proved. We denote the central point of the k lines (where k is even), obtained from our n lines by discarding the $n - k$ lines with numbers $i, j, \ldots, r$, by $z_{ij\ldots r}$, and we denote the central circle of the l lines (where l is odd), obtained from our lines by discarding the $n - l$ lines with numbers $i, j, \ldots, s$, by $S_{ij\ldots s}$. If n is *odd*, the problem amounts to showing that *the n central points z_1, z_2,..., z_n of all possible combinations of $n - 1$ of our lines lie on one circle*; to convince ourselves of this it is sufficient to show that *any four* of these points, for example z_1, z_2, z_3, and z_4, lie on one circle. But these points can be regarded as the points of intersection of the circles S_{14} and S_{12}, of S_{12} and S_{23}, of S_{23} and S_{34}, and of S_{34} and S_{14}; the second points of intersection of these pairs of circles, the points z_{124}, z_{123}, z_{234}, and z_{134}, lie on one circle, S_{1234}. If n is *even*, the problem amounts to showing that *the n central circles S_1, S_2,..., S_n of all possible combinations of $n - 1$ of our lines meet in one point*; here it is sufficient to verify that *any three* of these circles, for example S_1, S_2, and S_3, meet in one point.[23] But this follows by considering the four circles S_1, S_{134}, S_{234}, and S_2. From the fact that the points z_{14}, z_{1234}, z_{24}, and z_{12}, in which these circles intersect in pairs, lie on the one circle S_{124}, it follows that their second points of intersection, points z_{13}, z_{34}, z_{23}, and the point of intersection of S_2 and S_1, also lie on one circle; that is, the point of intersection of S_2 and S_1 also lies on the circle S_3, which passes through the points z_{13}, z_{34}, and z_{23}.

There are many more analogous theorems. We consider two lines in general position: on each of these lines we choose arbitrarily a point other than their point of intersection. We call the circle which passes through these two points and the point of intersection of the two given lines the **director circle**[24] of the two lines and the points given on them (Figure 9a). We consider, then, three lines in general position with points given on them; in this case *the three director circles, corresponding*

---

[23] It is not difficult to construct four circles, of which any three meet in one point, but which do not all have a common point. However, it is fairly simple to convince oneself that if the number of circles *exceeds four*, then from the fact that any three of the circles meet in one point it necessarily follows that all the circles have a common point.

[24] This has, of course, no connection with the director circle of a conic section (the locus of points from which the tangents to the conic are perpendicular).—TRANSL.

to the three pairs of lines which can be chosen from our three lines, meet in one point (Figure 9b); this point may be called the **director point** of the three lines. We now take in the plane four lines in general position, and on each of them we choose a point; in addition we need all these points to *lie on one circle* (or line). In this case the four director points of the four triads of lines which can be chosen from our four lines will always lie on one circle; this circle may be called the director circle of our four lines (Figure 9c).

Before proceeding further, we may note that this proposition is equivalent to the following. *If we take points on the sides of an arbitrary quadriangle* (possibly nonconvex or self-intersecting), *so that all these points lie on one circle* (or line), and *join them successively and describe the circumcircles of the four triangles thus formed, then the four points of intersection of adjacent* circles lie on one circle, the director circle of the sides of the quadrangle (Figure 9d).

By continuing our series of definitions above we can, in exactly the same way, associate with any odd number of lines in general position, on each of which a point is chosen (so that all the points lie on one circle or line), the director point of these lines, and with any even number of lines in general position (the points lying on them situated on one circle or line) the director circle; moreover, the director point of n lines is defined as the point of intersection of the n director circles of all possible combinations of $n - 1$ of the given lines, and the director circle of n lines is defined as the circle which passes through the n director points of all possible combinations of $n - 1$ of the given lines.

In order to prove the statement in the case of *three* lines, it is sufficient to take for the circles S_1, S_2, S_3, and S_4 of the original theorem the line 1, the director circle S_{13} of the lines 1 and 3, the director circle S_{23} of the lines 2 and 3, and the line 2 (Figure 9b). For the points z_1, z_2, z_3, and z_4 we take the point u_{13} (the point of intersection of the lines 1 and 3), the point u_3 chosen on the line 3, the point u_{23} (the point of intersection of the lines 2 and 3), and the "point" ∞, which lies on both the lines 1 and 2; these four points lie on one "circle or line" (in fact, line 3). The role of the points w_1, w_2, w_3, and w_4 is played by the point u_1 chosen on the line 1, the point u_{123} (the point of intersection of the circles S_{13} and S_{23}), the point u_2 chosen on the line 2, and the point u_{12} (the point of intersection of the lines 1 and 2); since these points must also lie on one circle, the point u_{123} will always lie also on the director circle S_{12} of the lines 1 and 2, which passes through the points u_1, u_2, and u_{12}.

To prove the statement in the case of *four* lines it is sufficient to take for the circles S_1, S_2, S_3, and S_4 (which figure in the condition of our original theorem) the circles S_{12}, S_{23}, S_{34}, and S_{41} where, for example, S_{12} is the director circle of the lines 1 and 2, and to take for the points z_1, z_2, z_3, and z_4 the points u_2, u_3, u_4, and u_1 chosen on the lines

2, 3, 4, and 1 (see Figure 9d; the points u_1, u_2, u_3, and u_4 lie on one circle or line by the hypothesis of the theorem). In this case the points u_{123}, u_{234}, u_{134}, and u_{124} are the second points of intersection of our circles where, say, u_{123} is the director point of the lines 1, 2, and 3; by virtue of the proof of the theorem stated above these points will lie on one circle S_{1234}, the director circle of the lines 1, 2, 3, and 4.

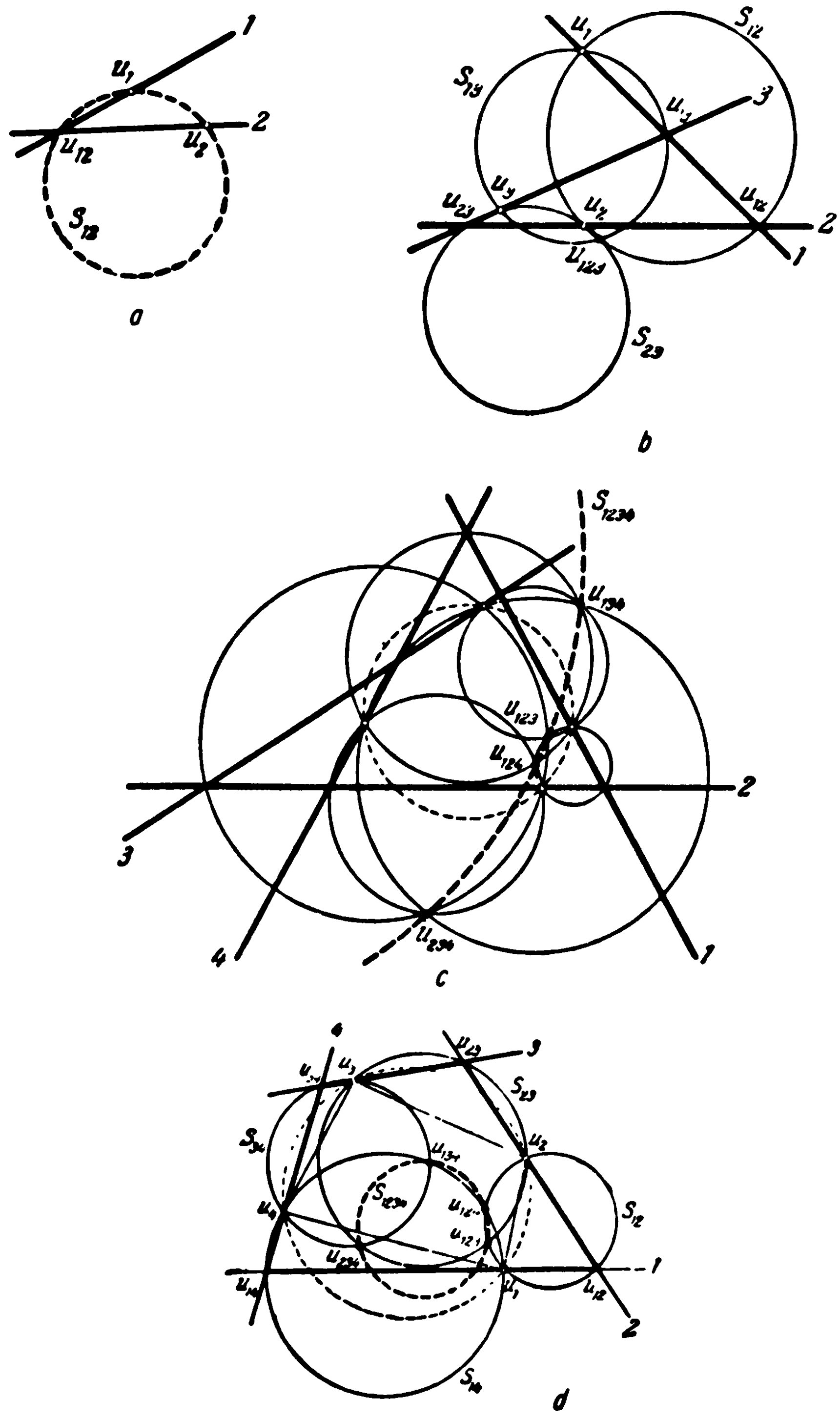


FIG. 9a, b, c, d.

Let us consider now n lines in general position 1, 2, 3,..., n, with points $u_1, u_2, u_3, \ldots, u_n$ on them, lying on one line or circle. Let us suppose that for any number of lines less than n our theorem is already proved; let us denote the director circle of the k lines (k even), obtained from the given n lines by discarding the $n - k$ lines with numbers $i, j, \ldots, r$, by $S_{ij\ldots r}$, and denote the director point of the l lines (l odd), obtained from the given lines by discarding the $n - l$ lines with numbers $i, j, \ldots, s$, by $u_{ij\ldots s}$. If $n \geqslant 6$ is *even*, the problem amounts to showing that *the n director points $u_1, u_2, \ldots, u_n$ of all possible combinations of $n - 1$ of the given lines lie on one circle*, for which it is sufficient to ensure that *any four* of these points, for example the points u_1, u_2, u_3, and u_4, lie on one circle. If $n \geqslant 5$ is *odd*, it is necessary to show that *the n director circles $S_1, S_2, \ldots, S_n$ of all possible combinations of $n - 1$ of our lines meet in one point*, and for this it is sufficient to ensure that *any three* of these circles, say S_1, S_2, and S_3, meet in one point. The proof of the last statements does not differ essentially from the reasoning used in the proof of the theorems of Clifford (for arbitrary n).

We now turn to some other ideas and facts in which the study of complex numbers can help us. We recall the equation of a circle S (Equation 14, Section 7):

$$Az\bar{z} + Bz - \bar{B}\bar{z} + C = 0, \qquad A, C \text{ purely imaginary}$$

Let a line l passing through the point O intersect S at the points z_1 and z_2; see Figure 10. We shall determine the product of the lengths

$$\{O, z_1\} \cdot \{O, z_2\}$$

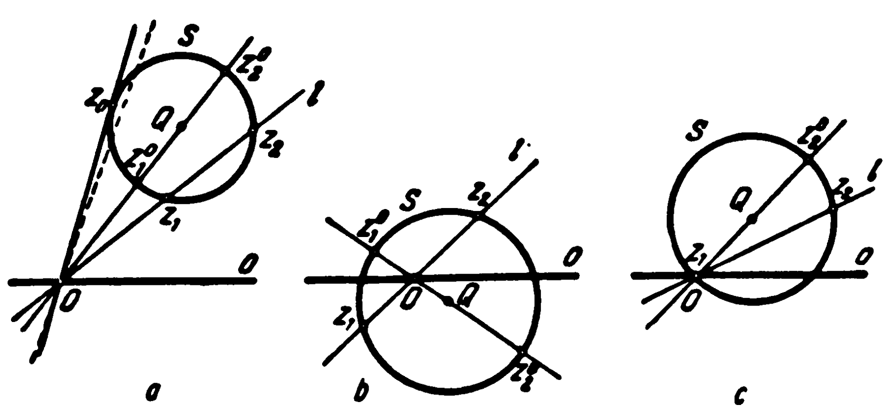


FIG. 10

where, as before, the braces emphasize that it is a question of *oriented* lengths of segments; that is, the product $\{O, z_1\} \cdot \{O, z_2\}$ is counted as positive if the directions of the **segments** $\overline{Oz_1}$ and $\overline{Oz_2}$ (from O to z_1 and from O to z_2, respectively) coincide and the points z_1 and z_2 lie on the same side of O, and negative if the directions of these segments are opposite (the points z_1 and z_2 lie on opposite sides of O).

Since the points z_1 and z_2 lie on the circle S,

$$Az_1\bar{z}_1 + Bz_1 - \bar{B}\bar{z}_1 + C = 0 \tag{16}$$

$$Az_2\bar{z}_2 + Bz_2 - \bar{B}\bar{z}_2 + C = 0 \tag{17}$$

On the other hand, since these points lie on a line through the origin O, $\arg z_2 = \arg z_1$ and $\arg \bar{z}_2 = -\arg z_1$ (Figure 10a), or $\arg z_2 = \arg z_1 + \pi$ and $\arg \bar{z}_2 = -\arg z_1 - \pi$ (Figure 10b), and consequently the product $z_1\bar{z}_2$ is real in all cases:

$$z_1\bar{z}_2 = k, \qquad \bar{z}_1 z_2 = \overline{z_1\bar{z}_2} = \bar{k} = k \tag{18}$$

We now multiply Equation 16 by z_2 and Equation 17 by z_1 and use Equations 18; we obtain

$$Akz_1 + Bz_1z_2 - \bar{B}k + Cz_2 = 0 \tag{16a}$$

$$Akz_2 + Bz_1z_2 - \bar{B}k + Cz_1 = 0 \tag{17a}$$

Subtracting Equation 17a from Equation 16a, we have

$$Ak(z_1 - z_2) - C(z_1 - z_2) = 0$$

whence it follows that, if $z_1 \neq z_2$, so that $z_1 - z_2 \neq 0$, then

$$k = \frac{C}{A}$$

But the product

$$z_1\bar{z}_2 = k$$

exactly coincides with the product $\{O, z_1\} \cdot \{O, z_2\}$ of the lengths of the (oriented) segments $\overline{Oz_1}$ and $\overline{Oz_2}$. In fact, obviously

$$|k| = |z_1| \cdot |\bar{z}_2| = |z_1| \cdot |z_2|$$

is equal to the product $(O, z_1) \cdot (O, z_2)$ of the lengths of the (nonoriented) segments $\overline{Oz_1}$ and $\overline{Oz_2}$. On the other hand, the

number k is positive if the points z_1 and z_2 lie on the same side of O, and $\arg \bar{z}_2 = -\arg z_1$; the number k is negative if the points z_1 and z_2 lie on opposite sides of O, and $\arg \bar{z}_2 = -\arg z_1 - \pi$.

Thus we have finally

$$\{O, z_1\} \cdot \{O, z_2\} = k = \frac{C}{A} \tag{19}$$

We have deduced this relation on the assumption that the points z_1 and z_2 are different, but it is clear that if $z_1 = z_2 = z_0$ (see Figure 10a), then the product $\{O, z_1\} \cdot \{O, z_2\} = \{O, z_0\}^2$ is also equal to C/A; this follows from the fact that we can regard the quantity $\{O, z_0\}^2$ as the limit of the expression $\{O, z_1\} \cdot \{O, z_2\}$, where z_1 and z_2 are the points of intersection of S and the secant $[O, z_1]$, very close to the tangent $[O, z_0]$ (and tending towards $[O, z_0]$). On the other hand, if the circle S degenerates to a point ("circle of zero radius"), then the expression $k = C/A$ is equal to the square of the distance (O, S), since in this case there exists a unique "secant" $[Oz_1z_2]$ of the "circle" S, for which $\{O, z_1\} \cdot \{O, z_2\} = (O, S)^2$. If the circle S passes through the point O (Figure 10c), one of the points z_1 and z_2 coincides with O, and $\{O, z_1\} \cdot \{O, z_2\} = 0$; on the other hand, in this case $k = C/A = 0$, since in Equation 14 $C = 0$ (as the point $z = 0$ satisfies this equation).

The expression $k = C/A$ is called the **power of the circle** S (more precisely, *the power of the point* O *with respect to* S);[25] its geometrical meaning is given by Equation 19 (thus, *the product* $\{O, z_1\} \cdot \{O, z_2\}$ *does not depend on the choice of the secant* l *of the circle which passes through* O). If the point O lies outside the

---

[25] The ratio $C/A = k$ depends, not only on the circle S, but also on the system of (complex) coordinates in which the equation of S has the form of Equation 14. However, it is not difficult to see that in fact the quantity C/A depends only on the position of the origin O, and not on the direction of the real axis o. This follows from the fact that by a *rotation about* O, indicated by $z' = pz$ and $z = p'z'$, where $|p| = 1$ and $|p'| = |1/p| = 1$ (cf. Equation 4), the circle, Equation 14, goes into the circle $Ap'\bar{p}'z'\bar{z}' + Bp'z' - \bar{B}\bar{p}'\bar{z}' + C = 0$, which has the same power $C/Ap'\bar{p}' = C/A = k$ (since $p'\bar{p}' = 1$).

circle S, the power of O with respect to S can also be defined as *the square* $\{O, z_0\}^2$ *of the length of the segment of the tangent drawn from* O *to* S (Figure 10a); further, by taking the secant l through the center Q of the circle S and denoting the distance (O, Q) by d and the radius of S by r, we obtain the result that *the power of* O *with respect to* S *is in all cases* (in magnitude and sign) *equal to*

$$(d + r)(d - r) = d^2 - r^2$$

We can also define the *power of an arbitrary point* w *with respect to the circle* S as the product

$$\{w, z_1\} \cdot \{w, z_2\}$$

where z_1 and z_2 are the points of intersection of S and a line passing through w (Figure 11). If w lies outside S, the power of w

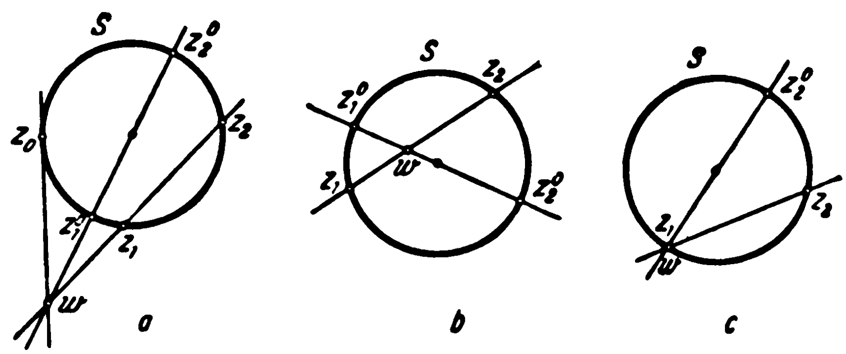


FIG. 11

with respect to S is equal to the square $\{w, z_0\}^2$ of the length of the segment of the tangent drawn from w to S (Figure 11a); if the radius of S is equal to r, and the distance from the center of S to w is equal to d, the power of w with respect to S is equal to $(d + r)(d - r) = d^2 - r^2$. Hence, in particular, it follows that *the power of* w *with respect to* S *is positive if* w *lies outside* S, *negative if* w *lies inside* S, *and zero if* w *lies on* S.

In order to define the numerical value of the power of the point w (where w is a definite complex number) with respect to the circle given by Equation 14, it is sufficient to introduce into the plane of the complex variable a new system of (complex) coordinates,

$$Z = z - w, \qquad z = Z + w$$

whose origin is the point w. In the new system of coordinates the circle, Equation 14, has the equation

$$A(Z + w)(\overline{Z + w}) + B(Z + w) - \bar{B}(\overline{Z + w}) + C = 0$$

or

$$AZ\bar{Z} + (A\bar{w} + B)Z - (-Aw - \bar{B})\bar{Z} + (Aw\bar{w} + Bw - \bar{B}\bar{w} + C) = 0$$

By using Equation 19 for the power of the origin O with respect to the circle (Equation 14) we conclude that the power of the point w with respect to the circle (Equation 14) is equal to

$$\frac{Aw\bar{w} + Bw - \bar{B}\bar{w} + C}{A} = w\bar{w} + \frac{B}{A}w - \frac{\bar{B}}{A}\bar{w} + \frac{C}{A} \quad (20)$$

In other words, *the power of the point w with respect to the circle* (Equation 14) *is the real number obtained by substituting w into the equation of the circle, normalized by taking the coefficient of $z\bar{z}$ to be equal to* 1 (that is, by dividing all the terms of Equation 14 by A).

This result enables us to solve a whole series of interesting problems on the determination of loci. Thus it follows that *the locus of points w, whose power with respect to the given circle* (Equation 14) *is a known quantity k*, is given by the equation

$$\frac{Aw\bar{w} + Bw - \bar{B}\bar{w} + C}{A} = k,$$

or

$$Aw\bar{w} + Bw - \bar{B}\bar{w} + (C - Ak) = 0$$

That is, *it is a circle* (moreover, as we see easily, it is a circle concentric with the original circle; see Footnote 22). On the other hand, if we have two circles S_1 and S_2 with equations

$$Az\bar{z} + B_1 z - \bar{B}_1\bar{z} + C_1 = 0$$

and

$$Az\bar{z} + B_2 z - \bar{B}_2\bar{z} + C_2 = 0$$

where for simplicity we take the coefficients of $z\bar{z}$ to be equal (this condition does not restrict the generality, since it can easily be satisfied by multiplying all the terms of one equation by a suitably chosen real number), then *the locus of points w,*

whose powers with respect to S_1 and S_2 are equal, is given by the equation

$$\frac{Aw\bar{w} + B_1w - \bar{B}_1\bar{w} + C_1}{A} = \frac{Aw\bar{w} + B_2w - \bar{B}_2\bar{w} + C_2}{A}$$

or

$$(B_1 - B_2)w - (\bar{B}_1 - \bar{B}_2)\bar{w} + (C_1 - C_2) = 0$$

That is, *it is a line q*; this line is called the **radical axis** of the circles S_1 and S_2. It is obvious that if the circles S_1 and S_2 have common points, then q passes through them (since each of these points has zero power with respect to S_1 and with respect to S_2); that is, it coincides with the common chord of S_1 and S_2 (Figure 12a). If S_1 and S_2 have no common points, then q can be characterized by the property that the segments of the tangents drawn from any point of this line to S_1 and to S_2 are equal (Figure 12b). Further, if we consider three circles S_1, S_2, and S_3 with equations

$$Az\bar{z} + B_1z - \bar{B}_1\bar{z} + C_1 = 0$$

$$Az\bar{z} + B_2z - \bar{B}_2\bar{z} + C_2 = 0$$

$$Az\bar{z} + B_3z - \bar{B}_3\bar{z} + C_3 = 0$$

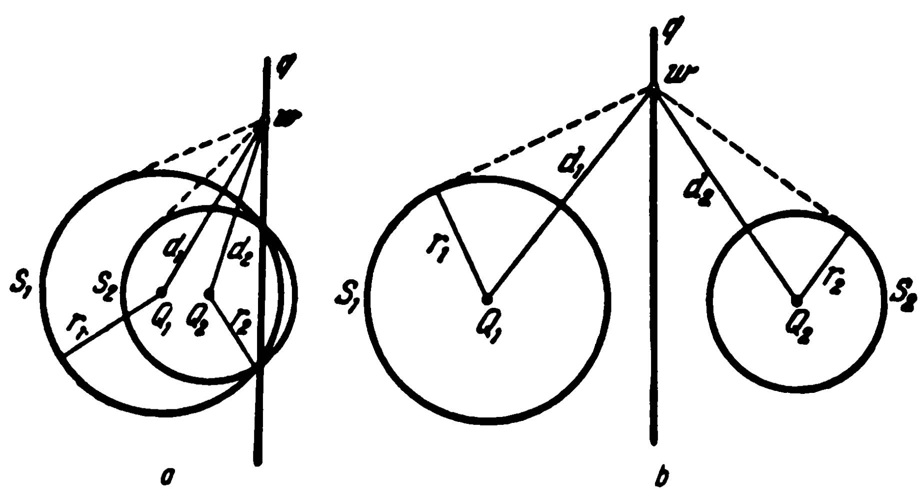


Fig. 12

then the radical axes of pairs of these circles are given by the equations

$$(B_1 - B_2)w - (\bar{B}_1 - \bar{B}_2)\bar{w} + (C_1 - C_2) = 0$$
$$(B_1 - B_3)w - (\bar{B}_1 - \bar{B}_3)\bar{w} + (C_1 - C_3) = 0$$
$$(B_2 - B_3)w - (\bar{B}_2 - \bar{B}_3)\bar{w} + (C_2 - C_3) = 0$$

But it follows from this that if the first two radical axes meet in some point Q, then the third radical axis passes through it (since the last of our three equations is the difference of the first two, so any solution of the first two equations satisfies it); thus, *the radical axes of three circles, taken in pairs, meet in one point*, called the **radical center** of the three circles (Figure 13), or *are parallel to each other*.

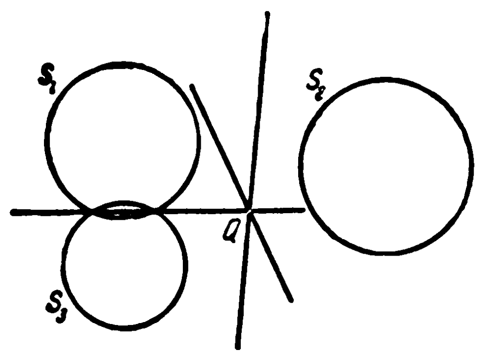


FIG. 13

The radical axis of two circles S_1 and S_2 may be defined as the locus of points for which the ratio of the powers with respect to S_1 and S_2 is equal to unity (or for which the difference of the powers with respect to S_1 and S_2 is equal to zero). We may also consider *the locus of points for which the difference of the powers with respect to S_1 and S_2 has a given value a* (if $a = 0$, we obtain the radical axis of the circles S_1 and S_2) and the *locus of points for which the ratio of the powers with respect to S_1 and S_2 has a given value α* (here we obtain the radical axis of S_1 and S_2 by putting $\alpha = 1$). It is obvious that the first of these two loci is given by the equation

$$\frac{Aw\bar{w} + B_1 w - \bar{B}_1\bar{w} + C_1}{A} - \frac{Aw\bar{w} + B_2 w - \bar{B}_2\bar{w} + C_2}{A} = a$$

or

$$(B_1 - B_2)w - (\bar{B}_1 - \bar{B}_2)\bar{w} + (C_1 - C_2 - Aa) = 0$$

That is, it is also *a line*. The second locus is given by the equation

$$\frac{Aw\bar{w} + B_1 w - \bar{B}_1\bar{w} + C_1}{A} : \frac{Aw\bar{w} + B_2 w - \bar{B}_2\bar{w} + C_2}{A} = \alpha$$

or

$$(1 - \alpha)Aw\bar{w} + (B_1 - \alpha B_2)w - (\bar{B}_1 - \alpha\bar{B}_2)\bar{w} + (C_1 - \alpha C_2) = 0$$

If $\alpha = 1$ it represents a *line* (the radical axis of S_1 and S_2), and if $\alpha \neq 1$ it represents a *circle.*

In particular, if the circles S_1 and S_2 reduce to points, we have the following. *The locus of points for which the difference of the squares of the distances from two given points S_1 and S_2 has a given value is a line*; *the locus of points for which the ratio of the squares of the distances from two given points S_1 and S_2 has a given value α is a line if $\alpha = 1$ and is a circle if $\alpha \neq 1$.* It is clear that in the latter case we can refer simply to *the ratio of the distances* (w, S_1) and (w, S_2) rather than the ratio of the squares of the distances. The circle which is the locus of points for which the ratio of the distances from two given points S_1 and S_2 has a fixed value α is often called the **circle of Apollonius** of these two points (from the outstanding Greek geometer Apollonius, who lived in in Asia Minor about the second century B.C.).

We now turn to the study of a triangle $\overline{a_1a_2a_3}$ (the square bracket across denotes that we are considering a *triangle* with vertices a_1, a_2, a_3 and not the product of three complex numbers a_1, a_2, and a_3; we shall use a similar notation later; cf. the notation for segments, earlier in this section). Let us suppose that $|a_1| = |a_2| = |a_3| = 1$; geometrically this means that all the vertices of the triangle lie on the unit circle $z\bar{z} = 1$ (thus we take the center of the circumcircle of the triangle to be the origin and the radius of the circle to be the unit of length); see Figure 14. In that case it is obvious that the point $a_1 + a_2 = h_3$ is a vertex of a rhombus $\overline{Oa_1h_3a_2}$, and so the lines $[Oh_3]$ and $[a_1a_2]$ are perpendicular (since they are diagonals of the rhombus); the

point $m_3 = h_3/2 = (a_1 + a_2)/2$ is the midpoint of the side $\overline{a_1a_2}$ of the triangle $\overline{a_1a_2a_3}$. Further, the point

$$h = a_1 + a_2 + a_3 \qquad (= h_3 + a_3)$$

is a vertex of a parallelogram $\overline{Oh_3ha_3}$. In other words, the line $[a_3h] \parallel [Oh_3] \perp [a_1a_2]$; that is, the line $[a_3h]$ is an *altitude* of the triangle $\overline{a_1a_2a_3}$, and the point b_3 where it meets the side $[a_1a_2]$ is the foot of the altitude. In exactly the same way it may be shown that the lines $[a_1h]$ and $[a_2h]$ are also altitudes of the triangle $\overline{a_1a_2a_3}$; hence $a_1 + a_2 + a_3 = h$ is the **orthocenter**, *the point of intersection of the altitudes of the triangle* $\overline{a_1a_2a_3}$.

From Figure 14 it is clear that the expression

$$(h_3\,, h) = (O, a_3) \qquad (= 1)$$

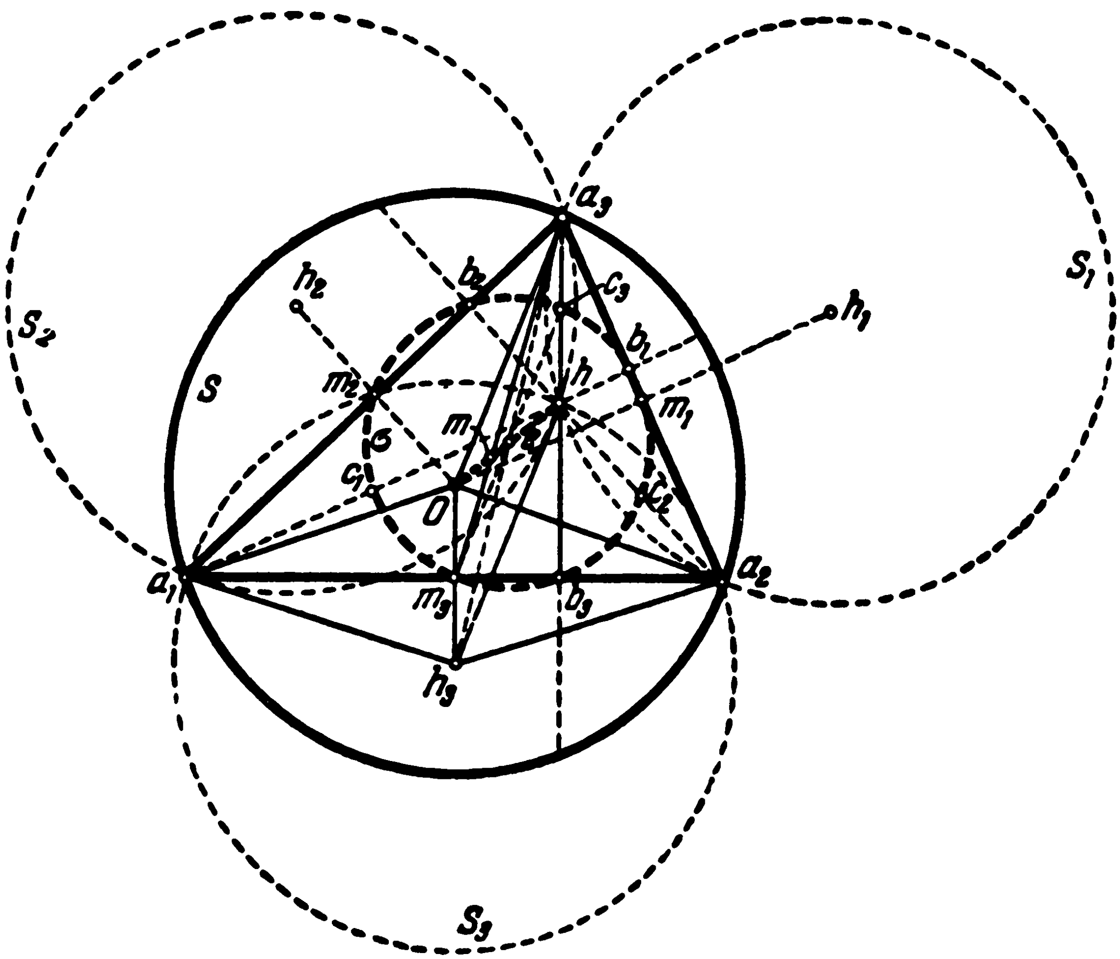


FIG. 14

which is *the distance from the orthocenter* h *of the triangle* $\overline{a_1a_2a_3}$ *to the point* h_3 (*the reflection of the circumcenter* O *in the side* $[a_1a_2]$), *is equal to the radius of the circumcircle* S *of the triangle*. Hence it follows that *the locus of the orthocenters of triangles* $\overline{a_1a_2a_3}$ *inscribed in* S, *of which two vertices* a_1 *and* a_2 *are fixed and the third moves on the circle* S, *is a circle equal to* S *with center at the point* $h_3 = a_1 + a_2$ (*the reflection of* O *in the side* $[a_1a_2]$). Further, if h_2 and h_1 are the reflections of O in the sides $[a_1a_3]$ and $[a_2a_3]$, then

$$[h_2, h] = (O, a_2) = 1, \qquad (h_1, h) = (O, a_1) = 1$$

Hence *the orthocenter* h *of an arbitrary triangle* $\overline{a_1a_2a_3}$ *coincides with the point of intersection of the circles* S_1, S_2, *and* S_3 (*equal to the circumcircle* S), *whose centers are* h_1, h_2, *and* h_3, *the reflections of the center* O *of the circle* S *in the sides of the triangle* (see Figure 14).

We consider now the point

$$e = \frac{h}{2} = \frac{a_1 + a_2 + a_3}{2}$$

It is clear that this is the point of intersection of the diagonals of the parallelogram $\overline{Oa_3hh_3}$; through this there passes also the line $[m_3c_3]$, which joins the midpoints of opposite sides of the parallelogram, where

$$c_3 = \frac{a_3 + h}{2} = \frac{a_1 + a_2}{2} + a_3$$

Here the point c_3 is the midpoint of the segment $\overline{a_3h}$ of the altitude $[a_3b_3]$ of the triangle. Furthermore,

$$(e, m_3) = (e, c_3) = \tfrac{1}{2}(O, a_3) = \tfrac{1}{2}$$

Thus, *the circle* σ *with center* e *and radius* $\frac{1}{2}$ *passes through the midpoint* m_3 *of the side* $\overline{a_1a_2}$ *of the triangle and through the midpoint* c_3 *of the segment* $\overline{a_3h}$ *of the altitude taken between the vertex and the orthocenter*. Similarly it may be shown that *this circle passes also through the midpoints* $m_1 = (a_2 + a_3)/2$ *and* $m_2 = (a_1 + a_3)/2$ *of the other two sides and through the midpoints* $c_1 = (a_2 + a_3)/2 + a_1$ *and* $c_2 = (a_1 + a_3)/2 + a_2$ *of the segments*

$\overline{a_1h}$ and $\overline{a_2h}$ of the other two altitudes. The circle σ was first considered by the great Swiss mathematican L. Euler (1707–1783); it is called the **Euler circle** of the triangle $\overline{a_1a_2a_3}$. Since the chords $[c_3b_3]$ and $[m_3b_3]$ of the circle σ are perpendicular, and $\overline{c_3m_3}$ is a diameter of this circle, the *Euler circle σ passes also through the foot b_3 of the altitude* $\overline{a_3b_3}$; similarly it may be shown that σ *passes also through the feet b_1 and b_2 of the other two altitudes $\overline{a_1b_1}$ and $\overline{a_2b_2}$ of the triangle*. (Thus, the circle σ passes through nine special points of the triangle $\overline{a_1a_2a_3}$, the points m_1, m_2, m_3, b_1, b_2, b_3, c_1, c_2, and c_3; hence it is often called the **nine-points circle** of the triangle.)

We note that *the point of intersection m of the medians* of the triangle $\overline{a_1a_2a_3}$, the **centroid**, or *center of gravity*, of the triangle divides the median $\overline{a_3m_3}$ in the ratio

$$(a_3, m):(m, m_3) = 2:1$$

Hence it is not difficult to see that it coincides with the point of intersection of the medians of the triangle $\overline{Oh_3a_3}$ (since $\overline{a_3m_3}$ is a median of this triangle too), and so m divides the median $\overline{Oe}$ of the triangle $\overline{Oh_3a_3}$ in the ratio

$$(O, m):(m, e) = 2:1$$

Thus, we see that *the point m lies on the line* $[Oe]$ and that $(O, m) = \frac{2}{3}(O, e)$, which is also $\frac{1}{3}(O, h)$; that is

$$m = \tfrac{2}{3}e = \frac{a_1 + a_2 + a_3}{3}$$

The line $[Oh]$ is called the **Euler line** of the triangle; it contains the center O of the circumcircle of the triangle $\overline{a_1a_2a_3}$, the point of intersection $m = (a_1 + a_2 + a_3)/3$ of the medians (the centroid of the triangle), the point of intersection $h = a_1 + a_2 + a_3$ of the altitudes (the orthocenter of the triangle), and the center $e = (a_1 + a_2 + a_3)/2$ of the Euler circle, where

$$(O, e) = \tfrac{1}{2}(O, h), \qquad (O, m) = \tfrac{1}{3}(O, h)$$

The theorems about the Euler circle can also be deduced without difficulty by direct calculation. We note first that, by virtue of Equations 6,

$$(e, m_3) = | e - m_3 | = \left| \frac{a_1 + a_2 + a_3}{2} - \frac{a_1 + a_2}{2} \right| = \left| \frac{a_3}{2} \right| = \frac{1}{2}$$

$$(e, c_3) = | e - c_3 | = \left| \frac{a_1 + a_2 + a_3}{2} - \left(\frac{a_1 + a_2}{2} + a_3 \right) \right|$$

$$= \left| \frac{-a_3}{2} \right| = \frac{1}{2}$$

Similarly it may be shown that

$$(e, m_1) = (e, m_2) = \tfrac{1}{2}, \qquad (e, c_1) = (e, c_2) = \tfrac{1}{2}$$

It is rather difficult to show that the circle σ with center e and radius $\frac{1}{2}$ passes through the points b_1 , b_2 , b_3 . To calculate the complex number b_3 we draw through a_3 the line $[a_3d_3] \parallel [a_2a_1]$; we denote the points of intersection of the lines $[a_3d_3]$ and $[a_3b_3]$ with the circle S by d_3 and f_3 . Since the arcs $\overset{\frown}{a_2a_3}$ and $\overset{\frown}{d_3a_1}$ of the circle S are equal, the angles which they subtend at the center are equal: $\angle a_2Oa_3 = \angle d_3Oa_1 = \alpha$. Hence,

$$\frac{a_3}{a_2} = \frac{a_1}{d_3} \qquad (= \cos \alpha + i \sin \alpha)$$

Thus we have

$$d_3 = \frac{a_1a_2}{a_3}$$

On the other hand, since $[d_3f_3]$ is a diameter of the circle S (because $[a_3d_3] \parallel [a_2a_1] \perp [a_3b_3]$), we have

$$f_3 = -d_3 = -\frac{a_1a_2}{a_3}$$

Further, since

$$(a_1 , h) = | a_1 - h | = | a_1 - (a_1 + a_2 + a_3)| = | a_2 + a_3 |$$

$$(a_1 , f_3) = | a_1 - f_3 | = \left| a_1 + \frac{a_1a_2}{a_3} \right| = \frac{| a_1 |}{| a_3 |} | a_2 + a_3 | = | a_2 + a_3 | ,$$

because $| a_1 | = | a_3 | = 1$, the triangle $\overline{a_1hf_3}$ is isosceles; therefore its altitude $[a_1b_3]$ coincides with a median, and b_3 is the midpoint of the segment $\overline{hf_3}$; whence it follows that

$$b_3 = \frac{h + f_3}{2} = \frac{a_1 + a_2 + a_3}{2} - \frac{a_1a_2}{2a_3}$$

Now it is easy to see that

$$(e, b_3) = | e - b_3 | = \left| \frac{a_1 + a_2 + a_3}{2} - \left(\frac{a_1 + a_2 + a_3}{2} - \frac{a_1 a_2}{2a_3} \right) \right|$$

$$= \left| \frac{a_1 a_2}{2a_3} \right| = \frac{| a_1 | \cdot | a_2 |}{2 \, | a_3 |} = \frac{1}{2}$$

Similarly we may show that

$$(e, b_1) = (e, b_2) = \tfrac{1}{2}$$

We proceed now to a *quadrangle* $\overline{a_1 a_2 a_3 a_4}$ inscribed in a circle S (Figure 15). As before, we take the center of this circle as the origin O and the radius of S as the unit of length. By analogy with the previous work we shall call the point

$$m = \frac{a_1 + a_2 + a_3 + a}{4}$$

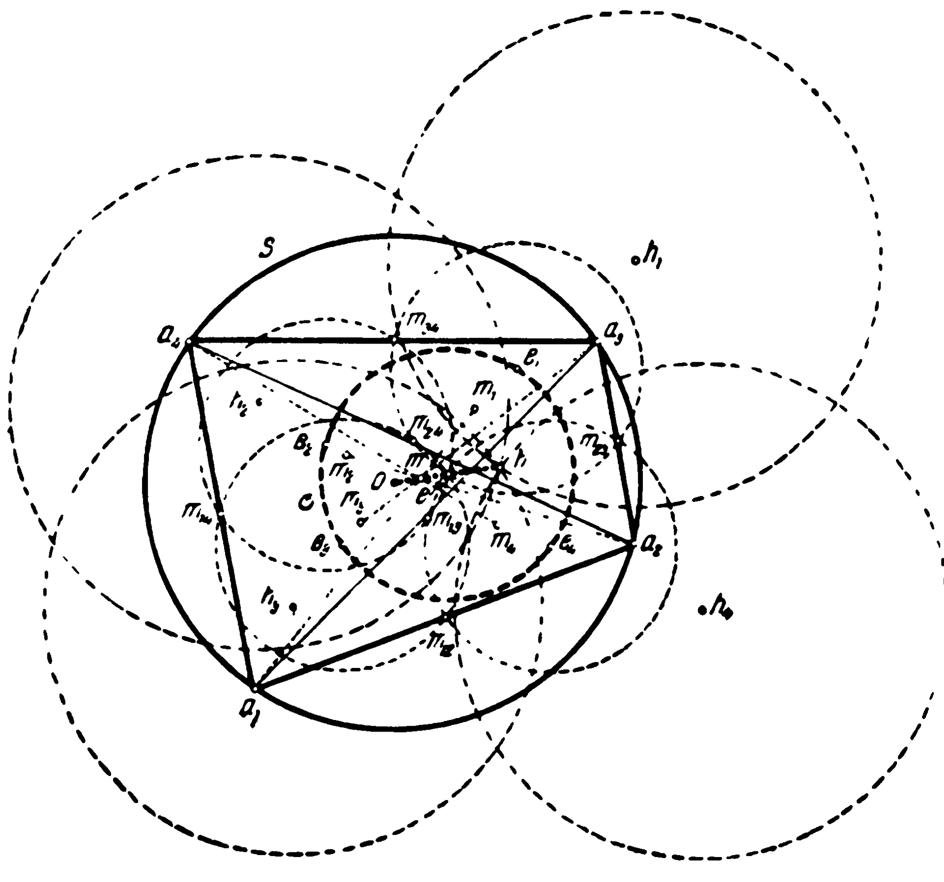


FIG. 15

the **centroid** (or center of gravity) of the quadrangle $\overline{a_1a_2a_3a_4}$, the point

$$h = a_1 + a_2 + a_3 + a_4$$

its **orthocenter**, and the circle σ of radius $\frac{1}{2}$ with center at the point

$$e = \frac{a_1 + a_2 + a_3 + a_4}{2}$$

the **Euler circle** of the quadrangle. Our next problem will consist in giving geometrical interpretations to all these expressions.

We consider the centroids m_4, m_3, m_2, and m_1, the orthocenters h_4, h_3, h_2, and h_1, and the centers of the Euler circles e_4, e_3, e_2, and e_1 of the triangles $\overline{a_1a_2a_3}$, $\overline{a_1a_2a_4}$, $\overline{a_1a_3a_4}$, and $\overline{a_2a_3a_4}$. We note first that

$$(h, h_4) = | h - h_4 |$$
$$= |(a_1 + a_2 + a_3 + a_4) - (a_1 + a_2 + a_3)| = | a_4 | = 1$$

and similarly

$$(h, h_1) = (h, h_2) = (h, h_3) = 1$$

Thus, *the four circles, equal to the circumcircle S of the quadrangle, whose centers coincide with the orthocenters of the triangles $\overline{a_1a_2a_3}$, $\overline{a_1a_2a_4}$, $\overline{a_1a_3a_4}$, and $\overline{a_2a_3a_4}$, meet in one point h*; we have called this point the orthocenter of the quadrangle $\overline{a_1a_2a_3a_4}$. Further,

$$(e, e_4) = | e - e_4 | = \left| \frac{a_1 + a_2 + a_3 + a_4}{2} - \frac{a_1 + a_2 + a_3}{2} \right|$$
$$= \left| \frac{a_4}{2} \right| = \frac{| a_4 |}{2} = \frac{1}{2}$$

and similarly

$$(e, e_1) = (e, e_2) = (e, e_3) = \tfrac{1}{2}$$

Hence it follows that *the Euler circles of the four triangles $\overline{a_1a_2a_3}$, $\overline{a_1a_2a_4}$, $\overline{a_1a_3a_4}$, and $\overline{a_2a_3a_4}$ meet in one point e, and the centers of these circles lie on one circle σ with center e and radius $\frac{1}{2}$*; we have called this circle the Euler circle of the quadrangle $\overline{a_1a_2a_3a_4}$.

Finally, from the fact that

$$a_4 - m = a_4 - \frac{a_1 + a_2 + a_3 + a_4}{4} = \frac{3a_4 - a_1 - a_2 - a_3}{4}$$

and

$$m - m_4 = \frac{a_1 + a_2 + a_3 + a_4}{4} - \frac{a_1 + a_2 + a_3}{3}$$

$$= \frac{3a_4 - a_1 - a_2 - a_3}{12}$$

which is to say that

$$a_4 - m = 3(m - m_4)$$

it follows that the point m lies on the segment $\overline{a_4 m_4}$ and divides it in the ratio

$$(a_4\,, m):(m, m_4) = 3:1$$

Similarly it can be shown that the point m lies on the segments $\overline{a_1 m_1}$, $\overline{a_2 m_2}$, *and* $\overline{a_3 m_3}$ and divides these segments in the ratio

$$(a_1\,, m):(m, m_1) = (a_2\,, m):(m, m_2) = (a_3\,, m):(m, m_3) = 3:1$$

In other words, *the four segments* $\overline{a_1 m_1}$, $\overline{a_2 m_2}$, $\overline{a_3 m_3}$, *and* $\overline{a_4 m_4}$, *joining each vertex of the quadrangle to the centroid of the triangle formed by the other three vertices, meet in one point* m and are divided by it in the ratio 3:1. We have called this point m the centroid of the quadrangle $\overline{a_1 a_2 a_3 a_4}$.

We may observe that *the points* O, $m = (a_1 + a_2 + a_3 + a_4)/4$, $e = (a_1 + a_2 + a_3 + a_4)/2$, *and* $h = a_1 + a_2 + a_3 + a_4$ *lie on one line*, where

$$(O, m) = \tfrac{1}{4}(O, h), \qquad (O, e) = \tfrac{1}{2}(O, h)$$

This line may be called the Euler line of the quadrangle.

This solves the problem of the geometrical definition of the centroid m, the orthocenter h, and the Euler circle σ of the quadrangle $\overline{a_1 a_2 a_3 a_4}$. However, we do not limit ourselves to these,

and we mention here some more theorems which characterize these points. First, it is easy to see that

$$m = \frac{a_1 + a_2 + a_3 + a_4}{4} = \frac{(a_1 + a_2)/2 + (a_3 + a_4)/2}{2}$$
$$= \frac{(a_1 + a_4)/2 + (a_3 + a_2)/2}{2} = \frac{(a_1 + a_3)/2 + (a_2 + a_4)/2}{2}$$

Hence it follows that the point m is the common midpoint of the segments $\overline{m_{12}m_{34}}$, $\overline{m_{14}m_{23}}$, and $\overline{m_{13}m_{24}}$, which join the midpoints m_{12} and m_{34} of the opposite sides $\overline{a_1a_2}$ and $\overline{a_3a_4}$ of the quadrangle, the midpoints m_{14} and m_{23} of the opposite sides $\overline{a_1a_4}$ and $\overline{a_2a_3}$, and the midpoints m_{13} and m_{24} of the diagonals $\overline{a_1a_3}$ and $\overline{a_2a_4}$; *the three segments joining the midpoints of opposite sides and the midpoints of the diagonals of the quadrangle* $\overline{a_1a_2a_3a_4}$ *meet in one point* m, *the centroid of the quadrangle*, and are bisected there; see Figure 16.

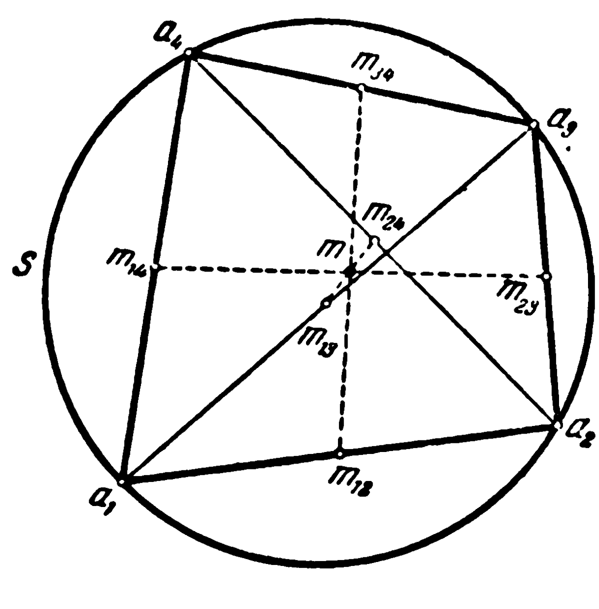


FIG. 16

The center e of the Euler circle σ of the quadrangle $\overline{a_1a_2a_3a_4}$ and the center O of the circumcircle are symmetrical about the centroid m. We may observe that

$$e = \frac{a_1 + a_2 + a_3 + a_4}{2} = \frac{(a_1 + a_2 + a_3) + a_4}{2} = \frac{h_4 + a_4}{2}$$

and similarly

$$e = \frac{h_1 + a_1}{2} = \frac{h_2 + a_2}{2} = \frac{h_3 + a_3}{2}$$

Hence it follows that the segments $\overline{a_1h_1}$, $\overline{a_2h_2}$, $\overline{a_3h_3}$, and $\overline{a_4h_4}$ all pass through the point e and are bisected there; *the four segments* $\overline{a_1h_1}$, $\overline{a_2h_2}$, $\overline{a_3h_3}$, *and* $\overline{a_4h_4}$, *which join each vertex of the quadrangle* $\overline{a_1a_2a_3a_4}$ *to the orthocenter of the triangle formed by the other three vertices, meet in one point* e, *the center of the Euler circle of the quadrangle, and are bisected there*; see Figure 17a.

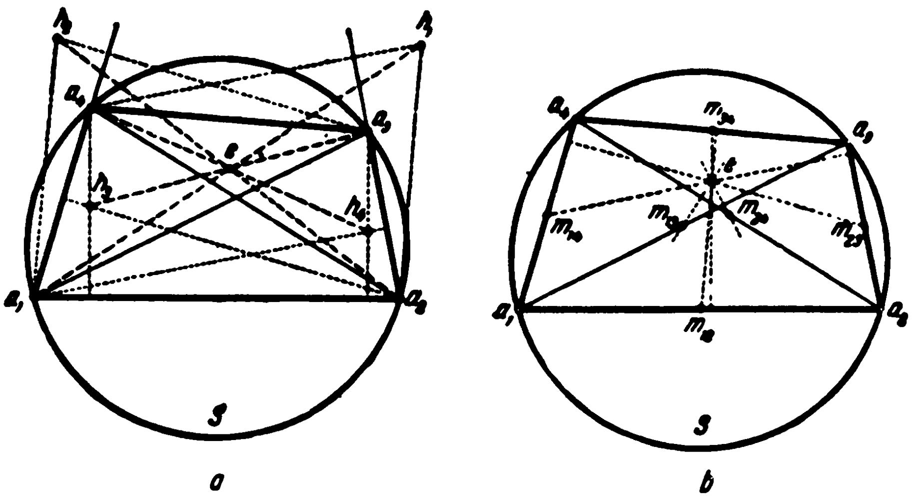


FIG. 17

We find the angle φ_{12} between the lines $[m_{12}e]$ and $[a_3a_4]$, where $m_{12} = (a_1 + a_2)/2$, as before, is the midpoint of the side $\overline{a_1a_2}$. From Equation 8 we have

$$\varphi_{12} = \arg \frac{e - m_{12}}{a_3 - a_4}$$

(we note that if e_{12} is the point of intersection of the lines $[m_{12}e]$ and $[a_3a_4]$, then we may write $\arg(e - e_{12}) = \arg(e - m_{12})$ and $\arg(a_3 - e_{12}) = \arg(a_3 - a_4)$). But

$$e - m_{12} = \frac{a_1 + a_2 + a_3 + a_4}{2} - \frac{a_1 + a_2}{2} = \frac{a_3 + a_4}{2}$$

and by the rule for the division of complex numbers we obtain

$$\frac{e - m_{12}}{a_3 - a_4} = \frac{(a_3 + a_4)(\bar{a}_3 - \bar{a}_4)}{2(a_3 - a_4)(\bar{a}_3 - \bar{a}_4)} = \frac{a_3\bar{a}_3 + a_4\bar{a}_3 - a_3\bar{a}_4 - a_4\bar{a}_4}{2(a_3\bar{a}_3 - a_4\bar{a}_3 - a_3\bar{a}_4 + a_4\bar{a}_4)}$$
$$= \frac{\bar{a}_3 a_4 - a_3\bar{a}_4}{2(2 - \bar{a}_3 a_4 - a_3\bar{a}_4)}$$

remembering that $a_3\bar{a}_3 = a_4\bar{a}_4 = 1$. Hence it follows that this ratio is a *purely imaginary* number (because in the last fraction the denominator is real and the numerator is the difference of two conjugate numbers), and so

$$\varphi_{12} = \arg \frac{\bar{a}_3 a_4 - a_3\bar{a}_4}{2(2 - \bar{a}_3 a_4 - a_3\bar{a}_4)} = \pm\frac{\pi}{2}, \qquad [m_{12}e] \perp [a_3a_4]$$

In exactly the same way it can be shown that

$$[m_{23}e] \perp [a_4a_1], \qquad [m_{34}e] \perp [a_1a_2], \qquad [m_{41}e] \perp [a_2a_3],$$
$$[m_{13}e] \perp [a_2a_4], \qquad [m_{24}e] \perp [a_1a_3]$$

Thus we find that *the six perpendiculars dropped from the midpoints of the six sides and diagonals of the quadrangle* $\overline{a_1a_2a_3a_4}$ *to the opposite sides or second diagonal respectively, meet in one point* e, *the center of the Euler circle of the quadrangle*; see Figure 17b.

Finally, *the orthocenter* h *of the quadrangle* $\overline{a_1a_2a_3a_4}$ *and the center* O *of the circumcircle are symmetrical about the center* e *of the Euler circle.*

There is no difficulty in extending the majority of these results to an arbitrary polygon inscribed in a circle S. Let us consider, for example, a *pentagon* $\overline{a_1a_2a_3a_4a_5}$. As above, we take the circle S circumscribed about it as the unit circle $z\bar{z} = 1$ in the plane of the complex variable; see Figure 18. Let m_1, m_2, m_3, m_4, m_5, and h_1, h_2, h_3, h_4, h_5, and e_1, e_2, e_3, e_4, e_5 be the centroids, orthocenters, and centers of the Euler circles of the quadrangles $\overline{a_2a_3a_4a_5}$, $\overline{a_1a_3a_4a_5}$, $\overline{a_1a_2a_4a_5}$, $\overline{a_1a_2a_3a_5}$, and $\overline{a_1a_2a_3a_4}$. We call the points

$$m = \frac{a_1 + a_2 + a_3 + a_4 + a_5}{5}, \qquad h = a_1 + a_2 + a_3 + a_4 + a_5,$$
$$e = \frac{a_1 + a_2 + a_3 + a_4 + a_5}{2}$$

respectively the **centroid**, the **orthocenter**, and the center of the **Euler circle** of the pentagon $\overline{a_1a_2a_3a_4a_5}$; we take the radius of the Euler circle to be $\frac{1}{2}$.

Exactly as done above it can be shown that the five circles, equal to the circumcircle S of the pentagon $\overline{a_1a_2a_3a_4a_5}$ and having as centers the orthocenters h_1 , h_2 , h_3 , h_4 , and h_5 of the quadrangles $\overline{a_2a_3a_4a_5}$, $\overline{a_1a_3a_4a_5}$, $\overline{a_1a_2a_4a_5}$, $\overline{a_1a_2a_3a_5}$, and $\overline{a_1a_2a_3a_4}$, meet in one point h (*the geometrical definition of the orthocenter of the pentagon*). The five Euler circles σ_1 , σ_2 , σ_3 , σ_4 , σ_5 of these five quadrangles meet in one point e, and their centers lie on one circle σ with center e and radius $\frac{1}{2}$ (*the geometrical definition of the Euler circle of the pentagon*). Finally, the five segments $\overline{a_1m_1}$,

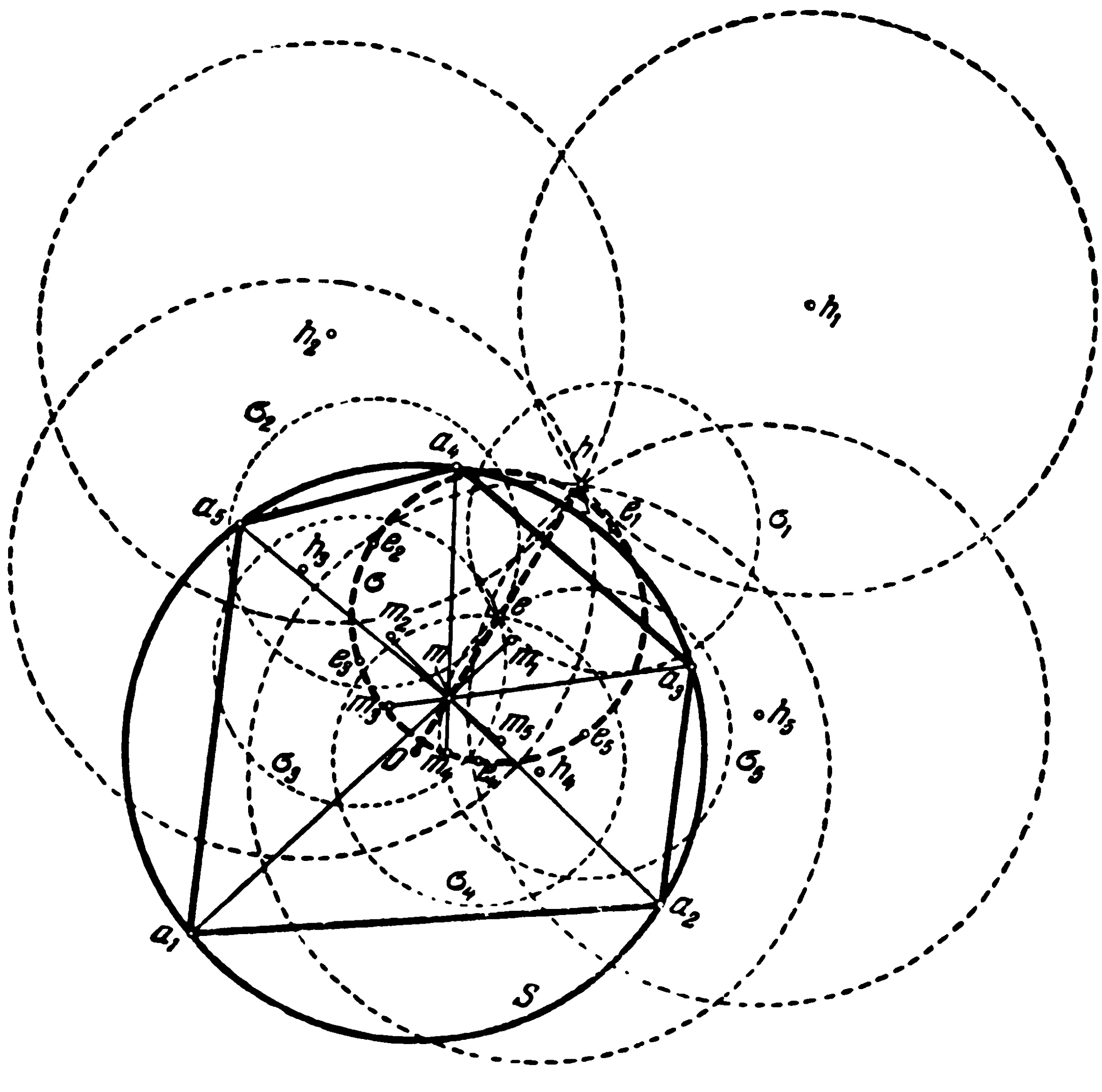


Fig. 18

$\overline{a_2m_2}$, $\overline{a_3m_3}$, $\overline{a_4m_4}$, and $\overline{a_5m_5}$, which join the vertices of the pentagon $\overline{a_1a_2a_3a_4a_5}$ to the centroids of the quadrangles formed by the other four vertices, meet in one point m and are divided there in the ratio 4:1, as reckoned from the vertices (*the geometrical definition of the centroid of the pentagon*). It is clear that the points O, m, e, and h lie on one line, where

$$(O, m) = \tfrac{1}{5}(O, h), \qquad (O, e) = \tfrac{1}{2}(O, h)$$

This line is naturally called the **Euler line** of the pentagon. It is not difficult to see also that the orthocenter h of the pentagon and the center O of the circumcircle are symmetrical about the center e of the Euler circle.

It is easy to see also that the ten segments, which join the midpoint of each side and each diagonal of the pentagon to the centroid of the triangle formed by the three vertices that do not lie on this side or diagonal, meet in one point m, the centroid of the pentagon, and are divided there in the ratio 3:2 as reckoned from the midpoints of the sides or diagonals. Further, the five segments, which join each vertex of the pentagon to the orthocenter of the quadrangle formed by the other four vertices, meet in one point, the center of the Euler circle of the pentagon, and are bisected there. Further, the ten perpendiculars, dropped from the centers of the Euler circles of the triangles formed by any three vertices of the pentagon to the segments joining the other two vertices (on sides or diagonals of the pentagon), meet in one point e, the center of the Euler circle of the pentagon.

Similarly, if we regard the idea of centroid, orthocenter, and Euler circle as having been already defined for all polygons inscribed in a circle, such that the number of sides is less than a given number n, then the **orthocenter** h of an n-gon $\overline{a_1a_2 \cdots a_n}$, incribed in a circle S may be defined as *the point of intersection of the n circles equal to S whose centers are the orthocenters of the $(n-1)$-gons formed by $n-1$ vertices of the n-gon.* The **Euler circle** σ of the n-gon is defined as the *circle which passes through the centers of the Euler circles $\sigma_1, \sigma_2, \ldots, \sigma_n$ of these n $(n-1)$-gons,* the center e of the circle σ being the point of intersection of the n circles $\sigma_1, \sigma_2, \ldots, \sigma_n$. Finally, the **centroid** m of an n-gon

inscribed in a circle S may be defined as *the point of intersection of the n segments which join each vertex of the n-gon to the centroid of the* $(n - 1)$*-gon formed by all the remaining vertices*, all these segments being divided at the point m in the ratio $(n - 1):1$ as reckoned from the vertices of the n-gon. *The points h, e, and m defined in this way lie on one line* passing also through O, where $(O, m) = \frac{1}{n}(O, h)$ and $(O, e) = \frac{1}{2}(O, h)$, the **Euler line** of the n-gon.

We note that all the segments which join the centroid of a k-gon, formed by k vertices of the n-gon $\overline{a_1 a_2 \cdots a_n}$, to the centroid of the $(n - k)$-gon, formed by the remaining $n - k$ vertices, pass through the point m, the centroid of the n-gon, and are divided there in the ratio $(n - k):k$ as reckoned from the centroid of the k-gon. Further, the n segments which join each vertex of the n-gon to the orthocenter of the $(n - 1)$-gon formed by the remaining $n - 1$ vertices meet in one point e, the center of the Euler circle of the n-gon, and are bisected there. The last theorem can also be generalized as follows. All the segments which join the orthocenter of a k-gon, formed by any k vertices of our n-gon, to the orthocenter of the $(n - k)$-gon, formed by the remaining $n - k$ vertices, pass through the point e and are bisected there. (Here the "orthocenter" of a chord $\overline{a_i a_j}$ of the circle S is understood to be the point $h_{i,j} = a_i + a_j$, the reflection of the center O of the circle S in this chord.) Moreover, the $n(n - 1)/2$ perpendiculars dropped from the centers of the Euler circles of the $(n - 2)$-gons formed by any $n - 2$ vertices of the n-gon to the segments which join the remaining two vertices (on a side or diagonal of the n-gon) meet in one point e, the center of the Euler circle of the n-gon. The number of similar theorems could be increased.

We now find the feet u_1, u_2, and u_3 of the perpendiculars dropped from any point u of the unit circle S in the complex plane to the sides $[a_2 a_3]$, $[a_3 a_1]$, and $[a_1 a_2]$ of a triangle $\overline{a_1 a_2 a_3}$ inscribed in the circle S; see Figure 19. On p. 53 it was shown that the foot of the perpendicular dropped from the point a_3 of the circle S to the chord $\overline{a_1 a_2}$ of the circle is expressed by the number

$$b_3 = \frac{1}{2}\left(a_1 + a_2 + a_3 - \frac{a_1 a_2}{a_3}\right)$$

Hence it follows that

$$u_3 = \frac{1}{2}\left(a_1 + a_2 + u - \frac{a_1 a_2}{u}\right)$$

$$u_1 = \frac{1}{2}\left(a_2 + a_3 + u - \frac{a_2 a_3}{u}\right)$$

$$u_2 = \frac{1}{2}\left(a_3 + a_1 + u - \frac{a_3 a_1}{u}\right)$$

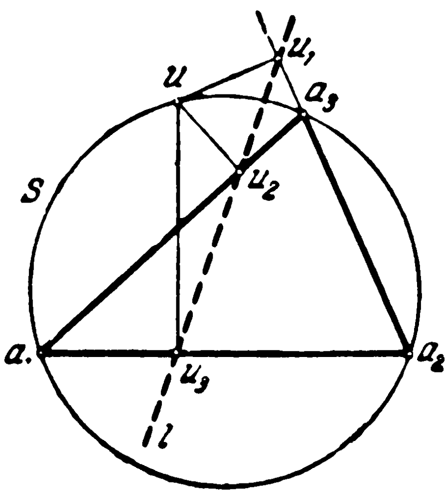


FIG. 19

We now observe that

$$V(u_1, u_2, u_3) = (u_1 - u_3) : (u_2 - u_3)$$

$$= \frac{1}{2}\left(a_3 - a_1 - \frac{a_2 a_3}{u} + \frac{a_1 a_2}{u}\right) : \frac{1}{2}\left(a_3 - a_2 - \frac{a_3 a_1}{u} + \frac{a_1 a_2}{u}\right)$$

$$= \left[(a_3 - a_1)\left(1 - \frac{a_2}{u}\right)\right] : \left[(a_3 - a_2)\left(1 - \frac{a_1}{u}\right)\right]$$

$$= \frac{(a_3 - a_1)(u - a_2)}{u} : \frac{(a_3 - a_2)(u - a_1)}{u}$$

$$= \frac{a_3 - a_1}{u - a_1} : \frac{a_3 - a_2}{u - a_2} = W(a_3, u, a_1, a_2)$$

But since the points a_3, u, a_1, and a_2 lie on one circle S, the cross-ratio $W(a_3, u, a_1, a_2)$ is real; therefore the ratio $V(u_1, u_2, u_3)$ also is real, and so *the three points* u_1, u_2, and u_3 *lie on one line*. This line is called the **Simson line** *of the point u with respect to the triangle* $\overline{a_1 a_2 a_3}$, after the English mathematician R. Simson (1687–1768), who established this fact.[25a]

[25a] This theorem was proved earlier by the famous English mathematician J. Wallis (1616–1703), so the traditional name "Simson line" contradicts historical accuracy.

We now derive the *equation of the Simson line l*. We start from Equation 10a, the equation of a line passing through two points z_1 and z_2; we normalize this equation by dividing all its terms by the coefficient of z:

$$z - \frac{z_1 - z_2}{\bar{z}_1 - \bar{z}_2}\bar{z} + \frac{z_1\bar{z}_2 - \bar{z}_1 z_2}{\bar{z}_1 - \bar{z}_2} = 0 \tag{21}$$

By putting $z_1 \equiv u_1$ and $z_2 \equiv u_3$ we obtain the following expression for the coefficient of $\bar{z}$:

$$\begin{aligned}(u_1 - u_3)/(\bar{u}_1 - \bar{u}_3) &= \frac{(a_3 - a_1)(u - a_2)}{u} : \frac{(\bar{a}_3 - \bar{a}_1)(\bar{u} - \bar{a}_2)}{\bar{u}} \\ &= \frac{(a_3 - a_1)(u - a_2)}{u} : \frac{(1/a_3 - 1/a_1)(1/u - 1/a_2)}{1/u} \\ &= \frac{(a_3 - a_1)(u - a_2)}{u} : \frac{(a_1 - a_3)(a_2 - u)}{a_1a_2a_3} = \frac{a_1a_2a_3}{u}\end{aligned}$$

We may note that, because $a_1\bar{a}_1 = 1$, we have $\bar{a}_1 = 1/a_1$ and, similarly, $\bar{a}_2 = 1/a_2$, $\bar{a}_3 = 1/a_3$, and $\bar{u} = 1/u$. Now, in order to determine the constant term C of Equation 21 it is sufficient to substitute into this equation $(z_1 - z_2)/(\bar{z}_1 - \bar{z}_2) = (a_1a_2a_3)/u$ and $z = u_1 = \frac{1}{2}(a_2 + a_3 + u - a_2a_3/u)$; then we obtain

$$\frac{1}{2}\left(a_2 + a_3 + u - \frac{a_2a_3}{u}\right) - \frac{1}{2}\left(\frac{a_1a_2a_3}{u}\right)\left(\bar{a}_2 + \bar{a}_3 + \bar{u} - \frac{\bar{a}_2\bar{a}_3}{\bar{u}}\right) + C = 0$$

whence, because $\bar{a}_2 = 1/a_2$, $\bar{a}_3 = 1/a_3$, $\bar{u} = 1/u$, and $\bar{a}_1 = 1/a_1$, we have

$$C = -\frac{1}{2}(a_1 + a_2 + a_3 + u) + \frac{1}{2}\left(\frac{a_1a_2a_3}{u}\right)(\bar{a}_1 + \bar{a}_2 + \bar{a}_3 + \bar{u})$$

Thus we finally arrive at the following equation:

$$\begin{aligned}z - \frac{a_1a_2a_3}{u}\bar{z} &- \frac{1}{2}(a_1 + a_2 + a_3 + u) \\ &+ \frac{1}{2}\left(\frac{a_1a_2a_3}{u}\right)(\bar{a}_1 + \bar{a}_2 + \bar{a}_3 + \bar{u}) = 0\end{aligned} \tag{22}$$

From Equation 22 we see immediately that the Simson line of the point u with respect to the triangle $\overline{a_1a_2a_3}$ passes through the point $z = (a_1 + a_2 + a_3 + u)/2$; if we agree to write a_4 instead of u, we obtain the result that *the Simson line of the vertex a_4 of a quadrangle $\overline{a_1a_2a_3a_4}$ inscribed in a circle S, with respect to the triangle $\overline{a_1a_2a_3}$ formed by the other three vertices, passes through the center $e = (a_1 + a_2 + a_3 + a_4)/2$*

of the Euler circle of the quadrangle. Hence follows one more definition of the center of the Euler circle of a quadrangle: *the four Simson lines of the four vertices of a quadrangle* $\overline{a_1a_2a_3a_4}$ *inscribed in a circle* S, *with respect to the triangles formed by the other three vertices, meet in one point* e, *the center of the Euler circle of the quadrangle* $\overline{a_1a_2a_3a_4}$; see Figure 20.

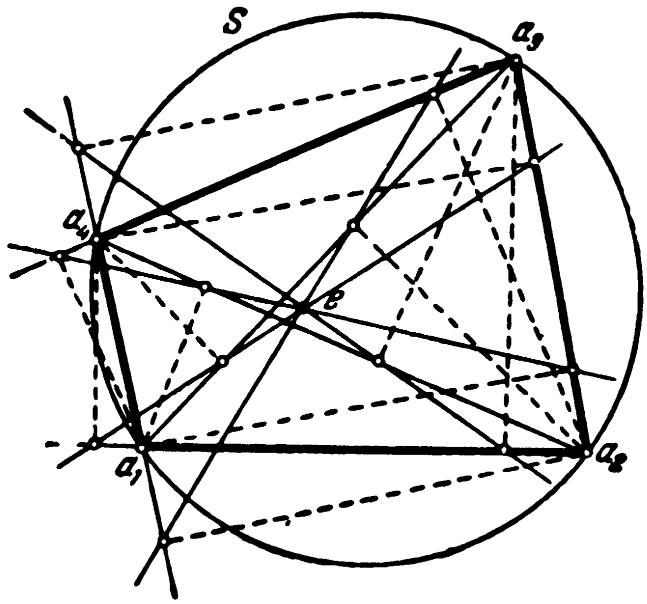


FIG. 20

It is not difficult to calculate that the foot of the perpendicular dropped from the point u to the line given by Equation 22 has the form

$$\frac{a_1 + a_2 + a_3 + 3u}{4} - \frac{a_1a_2a_3}{4u}(\bar{a}_1 + \bar{a}_2 + \bar{a}_3 - \bar{u}) \tag{23}$$

From this formula it can be deduced by direct calculation that *if* u *is a point of the circle* S *circumscribing a quadrangle* $\overline{a_1a_2a_3a_4}$, *then the feet of the perpendiculars dropped from the point* u *to the Simson lines of this point with respect to the triangles* $\overline{a_1a_2a_3}$, $\overline{a_1a_2a_4}$, $\overline{a_1a_3a_4}$, *and* $\overline{a_2a_3a_4}$ *lie on one line* (Figure 21); this line is called the Simson line of the point u with respect to the *quadrangle* $\overline{a_1a_2a_3a_4}$.

Similarly, the feet of the perpendiculars dropped from a point u of the circle S circumscribing a pentagon $\overline{a_1a_2a_3a_4a_5}$ to the Simson lines of this point with respect to the five quadrangles $\overline{a_1a_2a_3a_4}$, $\overline{a_1a_2a_3a_5}$, $\overline{a_1a_2a_4a_5}$, $\overline{a_1a_3a_4a_5}$, and $\overline{a_2a_3a_4a_5}$ lie on one line, the Simson line of the point u with respect to the *pentagon* $\overline{a_1a_2a_3a_4a_5}$. If, finally, we define in a similar way the Simson line of a point u of a circle S, with respect to any polygon

inscribed in S and having less than n sides, and then consider an n-gon $\overline{a_1a_2 \ldots a_n}$ inscribed in S, the feet of the perpendiculars dropped from the point u to the n Simson lines of this point with respect to all the $(n-1)$-gons formed by any $n-1$ vertices of the n-gon will all lie on one line, the Simson line of the point u with respect to the n-gon

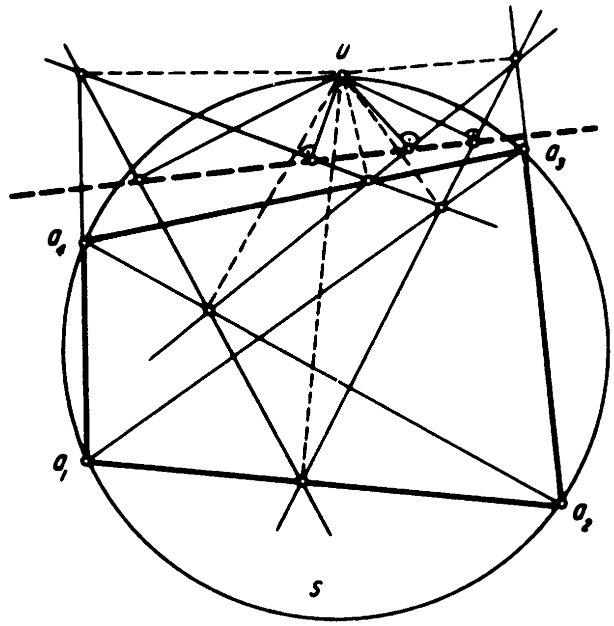


FIG. 21

$\overline{a_1a_2 \ldots a_n}$. If the circle S is taken as the unit circle in the plane of the complex variable, the equation of the Simson line of the point u with respect to the n-gon $\overline{a_1a_2 \ldots a_n}$ may be written as

$$z + (-1)^n \frac{s_n}{u^{n-2}} \bar{z} = \frac{(2^{n-2}-1)u^n + s_1u^{n-1} - s_2u^{n-2}}{2^{n-2}u^{n-1}} + \frac{s_3u^{n-3} + \ldots + (-1)^n s_{n-1}u + (-1)^n(2^{n-2}-1)s_n}{2^{n-2}u^{n-1}} \tag{24}$$

where $s_1 = a_1 + a_2 + a_3 + \cdots + a_n$, $s_2 = a_1a_2 + a_1a_3 + \cdots + a_{n-1}a_n$, $s_3 = a_1a_2a_3 + a_1a_2a_4 + \cdots + a_{n-2}a_{n-1}a_n, \ldots, s_n = a_1a_2 \ldots a_n$. We may note that Equation 22 can be written as:

$$z - \frac{a_1a_2a_3}{u}\bar{z} = \frac{u^3 + a_1u^2 + a_2u^2 + a_3u^2 - a_2a_3u - a_1a_3u - a_1a_2u - a_1a_2a_3}{2u^2}$$

or

$$z - \frac{s_3}{u}\bar{z} = \frac{(2-1)u^3 + s_1u^2 - s_2u - (2-1)s_3}{2u^2} \tag{22a}$$

whence it is obvious that it is equivalent to the particular case of Equation 24, which we obtain by putting $n = 3$.

Finally, we shall see that the definition of the Euler circle of a polygon $\overline{a_1a_2 \dots a_n}$ inscribed in a circle S can be considerably generalized. We associate with an n-gon $a_1a_2 \dots a_n$, where $n \geqslant 2$, inscribed in the unit circle S the n circles $\sigma^{(1)}, \sigma^{(2)}, \dots, \sigma^{(n)}$; the centers of these circles we take as the points $e^{(1)} = a_1 + a_2 + \cdots + a_n$, $e^{(2)} = (a_1 + a_2 + \cdots + a_n)/2$, $e^{(3)} = (a_1 + a_2 + \cdots + a_n)/3, \dots, e^{(n)} = (a_1 + a_2 + \cdots + a_n)/n$, and the radii we take as $1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}$. These circles may be called the first, second, third,..., and nth Euler circles of the n-gon; here the second Euler circle is the one which we simply called the "Euler circle" above, and the centers of the first and nth Euler circles coincide respectively with the orthocenter h and the centroid m of the n-gon. It is clear that for a segment $\overline{a_1a_2}$ the first Euler circle $\sigma^{(1)}$ is the reflection of the circle S in the line $[a_1a_2]$, and the Euler circle $\sigma^{(2)}$ has center at the midpoint of the segment and radius $\frac{1}{2}$; see Figure 22a. For a triangle $\overline{a_1a_2a_3}$ the first Euler

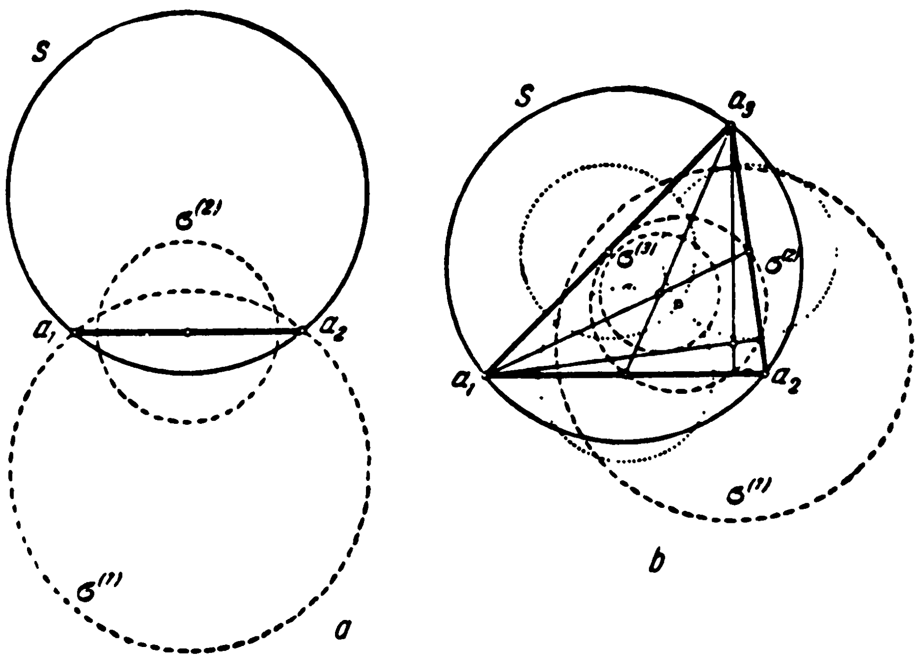


FIG. 22

circle $\sigma^{(1)}$ has center at the orthocenter of the triangle and radius 1; the second Euler circle $\sigma^{(2)}$ coincides with the nine-points circle; the third Euler circle $\sigma^{(3)}$ has center at the centroid of the triangle and radius $\frac{1}{3}$; see Figure 22b. Further, as is easily seen, *the centers of the* ith *Euler circles of the* n $(n - 1)$*-gons formed by any* $n - 1$ *vertices of an* n*-gon* $\overline{a_1a_2 \dots a_n}$ *inscribed in a circle* S *lie on the* ith *Euler circle of the* n*-gon* $\overline{a_1a_2 \dots a_n}$ (and these n circles meet in one point, the center $e^{(i)}$ of the

circle $\sigma^{(i)}$). Moreover, *the segment which joins the center of the* ith *Euler circle of a k-gon, formed by any k vertices of the polygon* $\overline{a_1 a_2 \dots a_n}$, *to the center of the* jth *Euler circle of the* $(n - k)$*-gon, formed by the remaining* $n - k$ *vertices, passes through the center* $e^{(i+j)}$ *of the* $(i + j)$th *Euler circle of the n-gon* (and is divided at this point in the ratio $j{:}i$ as reckoned from the center of the Euler circle of the k-gon). Finally, *the perpendiculars dropped from the* $n(n - 1)/2$ *centers of the* ith *Euler circles of all the* $(n - 2)$*-gons, formed by any* $n - 2$ *vertices of the n-gon* $\overline{a_1 a_2 \dots a_n}$, *to the segments formed by the remaining two vertices of* $\overline{a_1 a_2 \dots a_n}$ (on a side or diagonal of the n-gon) *meet in one point* $e^{(i)}$, the center of the ith Euler circle of the n-gon.

It is clear also that the centers of all the Euler circles of a polygon lie on one line, which passes through the center of the circumscribing circle, the Euler line of the polygon.

We now turn to the ratio of three points z_1, z_2, and z_3 of the plane,

$$V(z_1, z_2, z_3) = \frac{z_1 - z_3}{z_2 - z_3}$$

and the cross-ratio of four points z_1, z_2, z_3, and z_4,

$$W(z_1, z_2, z_3, z_4) = \frac{z_1 - z_3}{z_2 - z_3} : \frac{z_1 - z_4}{z_2 - z_4}$$

We shall explain how these quantities change as the order of the points is permuted.

Let us denote the original ratio

$$V(z_1, z_2, z_3) = \frac{z_1 - z_3}{z_2 - z_3}$$

by λ; then, obviously,

$$V(z_2, z_1, z_3) = \frac{z_2 - z_3}{z_1 - z_3} = \frac{1}{\lambda}$$

$$V(z_1, z_3, z_2) = \frac{z_1 - z_2}{z_3 - z_2} = \frac{(z_2 - z_3) - (z_1 - z_3)}{z_2 - z_3} = 1 - \lambda \tag{25}$$

The ratio of the three points is changed to its reciprocal if the first two points are interchanged, and to its complement with respect to unity if the last two points are interchanged. By applying these

rules successively we obtain

$$\begin{aligned} V(z_3, z_1, z_2) &= \frac{1}{1-\lambda} \\ V(z_2, z_3, z_1) &= 1 - \frac{1}{\lambda} = \frac{\lambda - 1}{\lambda} \qquad (25a) \\ V(z_3, z_2, z_1) &= 1 : \frac{\lambda - 1}{\lambda} = \frac{\lambda}{\lambda - 1} \end{aligned}$$

Thus, the ratio of three points of the plane (three complex numbers z_1, z_2, and z_3), depending on the order in which the points are taken, can have six values:

$$\lambda, \lambda_1 = \frac{1}{\lambda}, \qquad \lambda_2 = 1 - \lambda, \qquad \lambda_3 = \frac{1}{1-\lambda}$$

$$\lambda_4 = \frac{\lambda - 1}{\lambda}, \qquad \lambda_5 = \frac{\lambda}{\lambda - 1}$$

We now consider any triangle $\overline{z_1 z_2 z_3}$. With this triangle we can associate the complex number $\lambda = V(z_1, z_2, z_3) = (z_1 - z_3)/(z_2 - z_3)$; however, it is more accurate to talk of the six numbers λ, λ_1, λ_2, λ_3, λ_4, and λ_5 corresponding to our triangle. These six complex numbers are vertices of a hexagon $\overline{\lambda\lambda_3\lambda_1\lambda_2\lambda_5\lambda_4}$ (which may be degenerate or self-intersecting); this hexagon (Figure 23) is closely connected with the properties of original triangle[26] (thus, for example, all its vertices lie on one line if and only if all the vertices of the triangle $\overline{z_1 z_2 z_3}$ lie on one line). However, we limit ourselves here to answering only one question: *Under what circumstances is the hexagon* $\overline{\lambda\lambda_3\lambda_1\lambda_2\lambda_5\lambda_4}$ *degenerate, in the sense that two or more of its vertices coincide?*

Since any of the six possible values of the ratio of the three points can be taken as the "original value" λ, in order to solve

---

[26] It is clear that the whole hexagon $\overline{\lambda\lambda_3\lambda_1\lambda_2\lambda_5\lambda_4}$ is completely determined by a single vertex. To convince oneself of this it is sufficient to note that by virtue of the definition of the numbers λ, λ_1, λ_2, λ_3, λ_4, and λ_5 the lines $[O\lambda]$ and $[O\lambda_1]$, $[O\lambda_2]$ and $[O\lambda_3]$, and $[O\lambda_4]$ and $[O\lambda_5]$ are symmetrical about the axis o, that $(O, \lambda_1) = 1/(O, \lambda)$, $(O, \lambda_3) = 1/(O, \lambda_2)$, and $(O, \lambda_5) = 1/(O, \lambda_4)$, and that the segments $\overline{\lambda\lambda_2}$, $\overline{\lambda_1\lambda_4}$, and $\overline{\lambda_3\lambda_5}$ meet in one point $\frac{1}{2}$, which is their common midpoint (cf. Figure 23).

the given problem it is sufficient to explain under what circumstances the quantity λ is equal to one of the complex numbers λ_1, λ_2, λ_3, λ_4, and λ_5. But if

$$\lambda = \lambda_1 = \frac{1}{\lambda}$$

then $\lambda = (z_1 - z_3)/(z_2 - z_3) = 1$, which is impossible if the points z_1, z_2, and z_3 are different, or $\lambda = -1$. If

$$\lambda = \lambda_2 = 1 - \lambda$$

then

$$\lambda = \frac{1}{2} \qquad \text{and} \qquad \lambda_4 = \frac{\lambda - 1}{\lambda} = -1$$

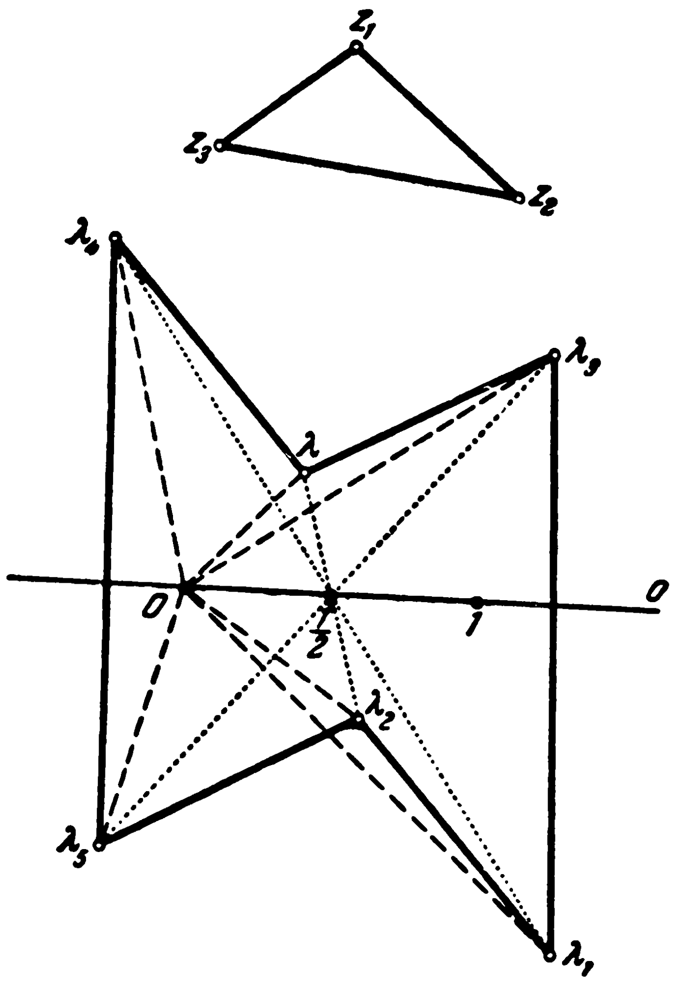


FIG. 23

If

$$\lambda = \lambda_5 = \frac{\lambda}{\lambda - 1}$$

then $\lambda = (z_1 - z_3)/(z_2 - z_3) = 0$, which also is impossible if z_1, z_2, and z_3 are different, or $\lambda = 2$ and $\lambda_2 = 1 - \lambda = -1$. On the other hand, if

$$\lambda = \lambda_3 = \frac{1}{1 - \lambda}$$

then λ satisfies the quadratic equation $\lambda^2 - \lambda + 1 = 0$, and so $\lambda = (1 + \sqrt{3}i)/2$ or $\lambda = (1 - \sqrt{3}i)/2$. If

$$\lambda = \lambda_4 = \frac{\lambda - 1}{\lambda}$$

then λ satisfies the same quadratic equation. We note that if $\lambda = (1 - \sqrt{3}i)/2$, then $\lambda_2 = 1 - \lambda = (1 + \sqrt{3}i)/2$.

Thus, we finally arrive at the conclusion that *the hexagon* $\overline{\lambda\lambda_3\lambda_1\lambda_2\lambda_5\lambda_4}$ (or $\overline{\lambda\lambda_1\lambda_2\lambda_3\lambda_4\lambda_5}$) *is degenerate only if one of the six values of the ratio of the three points* z_1, z_2, *and* z_3 *is equal to* -1 *or if one of the six values of the ratio of the three points is equal to* $(1 + \sqrt{3}i)/2$ ($= \cos 60° + i \sin 60°$). In the first case *one of the three points* z_1, z_2, *and* z_3 *is the midpoint of the segment joining*

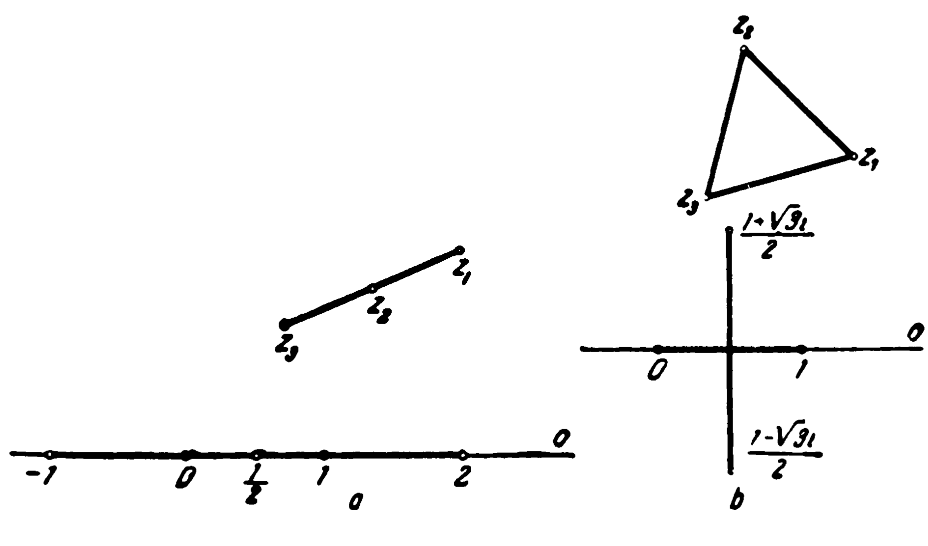


FIG. 24

the other two, and the hexagon $\overline{\lambda\lambda_1\lambda_2\lambda_3\lambda_4\lambda_5}$ degenerates into the 'triangle" $\overline{-1\tfrac{1}{2}2}$, the three "vertices" of which lie on one line o; see Figure 24a. In the second case *the triangle* $\overline{z_1z_2z_3}$ *is equilateral*, and the hexagon $\overline{\lambda\lambda_1\lambda_2\lambda_3\lambda_4\lambda_5}$ degenerates (see Figure 24b) into the segment

$$\overline{\frac{1+\sqrt{3}i}{2}\,\frac{1-\sqrt{3}i}{2}}$$

Similar deductions may be made concerning the cross-ratio

$$W(z_1, z_2, z_3, z_4) = \frac{z_1 - z_3}{z_2 - z_3} : \frac{z_1 - z_4}{z_2 - z_4}$$

of four points of the plane. From the definition of cross-ratio it follows that, if $W(z_1, z_2, z_3, z_4) = \lambda$, then

$$W(z_2, z_1, z_3, z_4) = \frac{z_2 - z_3}{z_1 - z_3} : \frac{z_2 - z_4}{z_1 - z_4} = \frac{1}{\lambda}$$

$$W(z_1, z_2, z_4, z_3) = \frac{z_1 - z_4}{z_2 - z_4} : \frac{z_1 - z_3}{z_2 - z_3} = \frac{1}{\lambda}$$

$$W(z_3, z_4, z_1, z_2) = \frac{z_3 - z_1}{z_4 - z_1} : \frac{z_3 - z_2}{z_4 - z_2} = \frac{z_1 - z_3}{z_1 - z_4} : \frac{z_2 - z_3}{z_2 - z_4} = \lambda$$

The cross-ratio of four points is changed to its reciprocal if the first two points or the last two points are interchanged, and does not change if the first pair of points is interchanged with the second pair. Further,

$$\begin{aligned} W(z_1, z_3, z_2, z_4) &= \frac{z_1 - z_2}{z_3 - z_2} : \frac{z_1 - z_4}{z_3 - z_4} = \frac{(z_2 - z_1)(z_3 - z_4)}{(z_2 - z_3)(z_1 - z_4)} \\ &= \frac{(z_2 - z_3)(z_1 - z_4) + (z_2 - z_4)(z_3 - z_1)}{(z_2 - z_3)(z_1 - z_4)} \\ &= 1 - \frac{z_1 - z_3}{z_2 - z_3} : \frac{z_1 - z_4}{z_2 - z_4} = 1 - \lambda \end{aligned}$$

The cross-ratio of four points is changed to its complement with respect to unity if the second and third points are interchanged. Successive application of these simple rules enables us to

conclude that *for all possible permutations of four points we obtain six different values of the cross-ratio*:

$$\begin{aligned}
W(z_1, z_2, z_3, z_4) &= W(z_2, z_1, z_4, z_3) = W(z_3, z_4, z_1, z_2) \\
&= W(z_4, z_3, z_2, z_1) = \lambda \\
W(z_1, z_2, z_4, z_3) &= W(z_2, z_1, z_3, z_4) = W(z_3, z_4, z_2, z_1) \\
&= W(z_4, z_3, z_1, z_2) = \frac{1}{\lambda} = \lambda_1 \\
W(z_1, z_3, z_2, z_4) &= W(z_2, z_4, z_1, z_3) = W(z_3, z_1, z_4, z_2) \\
&= W(z_4, z_2, z_3, z_1) = 1 - \lambda = \lambda_2 \\
W(z_1, z_3, z_4, z_2) &= W(z_2, z_4, z_3, z_1) = W(z_3, z_1, z_2, z_4) \\
&= W(z_4, z_2, z_1, z_3) = \frac{1}{1-\lambda} = \lambda_3 \\
W(z_1, z_4, z_2, z_3) &= W(z_2, z_3, z_1, z_4) = W(z_3, z_2, z_4, z_1) \\
&= W(z_4, z_1, z_3, z_2) = \frac{\lambda - 1}{\lambda} = \lambda_4 \\
W(z_1, z_4, z_3, z_2) &= W(z_2, z_3, z_4, z_1) = W(z_3, z_2, z_1, z_4) \\
&= W(z_4, z_1, z_2, z_3) = \frac{\lambda}{\lambda - 1} = \lambda_5
\end{aligned} \tag{26}$$

Thus, to each quadrangle $\overline{z_1z_2z_3z_4}$ there corresponds uniquely a hexagon $\overline{\lambda\lambda_1\lambda_2\lambda_3\lambda_4\lambda_5}$, whose properties are closely connected with the properties of the original quadrangle.[27]

We now explain under what circumstances the hexagon $\overline{\lambda\lambda_1\lambda_2\lambda_3\lambda_4\lambda_5}$ is degenerate, that is, has less than six distinct vertices. This question has already essentially been solved in the discussion of the ratio $V(z_1, z_2, z_3)$ of three points of the plane; in fact, the numbers λ_1, λ_2, λ_3, λ_4, and λ_5 are formed from the number λ in exactly the same way as described above. Thus, one example of a tetrad of points z_1, z_2, z_3, and z_4, whose cross-ratio, for various orders of the points, takes less than six different values, is a tetrad such that

$$W(z_1, z_2, z_3, z_4) = -1 \tag{27}$$

[27] See footnote 26.

Here, by changing the order of the points z_1, z_2, z_3, and z_4 in all possible ways, we obtain *three* different values of the cross-ratio of these points: -1, 2, and $\frac{1}{2}$. A tetrad of points z_1, z_2, z_3, and z_4 which satisfy Equation 27 is called a **harmonic tetrad**, and the corresponding quadrangle $\overline{z_1z_2z_3z_4}$ is called a **harmonic quadrangle**; see Figure 25a. Since in this case the cross-ratio

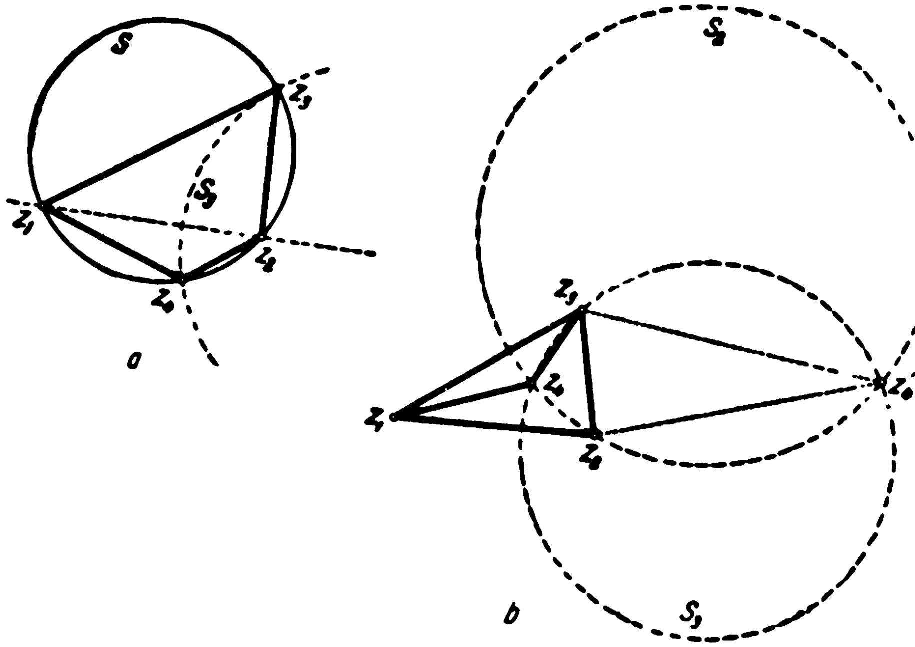


FIG. 25

$W(z_1, z_2, z_3, z_4)$ is real, *a circle can be circumscribed about a harmonic quadrangle* (cf. p. 32; this circle can degenerate into a line). On the other hand, from the equation

$$
\begin{aligned}
|W(z_1, z_2, z_3, z_4)| &= \left| \frac{z_1 - z_3}{z_2 - z_3} : \frac{z_1 - z_4}{z_2 - z_4} \right| \\
&= \frac{|z_1 - z_3| \cdot |z_2 - z_4|}{|z_2 - z_3| \cdot |z_1 - z_4|} \\
&= \frac{(z_1, z_3) \cdot (z_2, z_4)}{(z_2, z_3) \cdot (z_1, z_4)} = |-1| = 1
\end{aligned}
$$

it follows that

$$(z_1, z_3) \cdot (z_2, z_4) = (z_1, z_4) \cdot (z_2, z_3)$$

Thus, *the products of the lengths of opposite sides of a harmonic quadrangle are equal.* It is clear that these two conditions completely characterize a harmonic quadrangle: if a circle can be circumscribed about a quadrangle $\overline{z_1 z_2 z_3 z_4}$, then the cross-ratio $W(z_1, z_2, z_3, z_4)$ is real, and if, in addition,

$$\frac{(z_1, z_3)(z_2, z_4)}{(z_1, z_4)(z_2, z_3)} = \frac{|z_1 - z_3| \cdot |z_2 - z_4|}{|z_1 - z_4| \cdot |z_2 - z_3|} = 1$$

then $|W(z_1, z_2, z_3, z_4)| = 1$, whence it follows that $W = -1$ (since the cross-ratio of four *distinct* points cannot be equal to $+1$). Hence it follows, for example, that a square is a harmonic quadrangle.

For any three points z_1, z_2, and z_3 it is always possible to choose a fourth point z_4 such that the tetrad of points z_1, z_2, z_3, and z_4 is harmonic; z_4 *is the point of intersection of the circumcircle S of the triangle $\overline{z_1 z_2 z_3}$ and the circle of Apollonius S_3 of the points z_1 and z_2, which passes through the point z_3* (that is, the locus of points w such that $(w, z_1)/(w, z_2) = (z_3, z_1)/(z_3, z_2)$ or

$$(w, z_1) \cdot (z_3, z_2) = (w, z_2) \cdot (z_3, z_1);$$

see p. 49). Only in the case in which the point z_3 is the *midpoint* of the segment $\overline{z_1 z_2}$ and both "circles" S and S_3 degenerate into lines is their point of intersection z_4 nonexistent; in such case the role of z_4 is played by the point at infinity ∞, since then $W(z_1, z_2, z_3, \infty) = V(z_1, z_2, z_3) = (z_1 - z_3)/(z_2 - z_3) = -1$ (cf. p. 34).

Another example of a tetrad of points z_1, z_2, z_3, and z_4 which gives less than six values of the cross-ratio W is obtained by putting

$$W(z_1, z_2, z_3, z_4) = \cos 60^\circ + i \sin 60^\circ = \frac{1 + \sqrt{3}i}{2} \tag{28}$$

Here, by changing the order of the points in all possible ways, we have *two* different values of the cross-ratio:

$$\cos 60^\circ + i \sin 60^\circ$$

and

$$\cos(-60^\circ) + i \sin(-60^\circ) = \frac{1 - \sqrt{3}i}{2}$$

A tetrad of points z_1, z_2, z_3, and z_4 which satisfy Equation 28 is called an **equianharmonic tetrad** and the quadrangle $\overline{z_1z_2z_3z_4}$ is called an **equianharmonic quadrangle.** See Figure 25b. Since in this case

$$W(z_1, z_2, z_3, z_4) = \frac{(z_1, z_3) \cdot (z_2, z_4)}{(z_2, z_3) \cdot (z_1, z_4)}$$
$$= |\cos 60^\circ + i \sin 60^\circ| = 1$$

we have, as before, $(z_1, z_3) \cdot (z_2, z_4) = (z_1, z_4) \cdot (z_2, z_3)$. On the other hand, from the equations

$$|W(z_1, z_3, z_2, z_4)| = |\cos(-60^\circ) + i \sin(-60^\circ)| = 1$$
$$|W(z_1, z_4, z_2, z_3)| = |\cos 60^\circ + i \sin 60^\circ| = 1$$

it follows that $(z_1, z_2) \cdot (z_3, z_4) = (z_1, z_4) \cdot (z_2, z_3)$ and $(z_1, z_2) \cdot (z_3, z_4) = (z_1, z_3) \cdot (z_2, z_4)$. Thus, *the product of the lengths of (any two) opposite sides of an equianharmonic quadrangle is equal to the product of the lengths of the diagonals:*

$$(z_1, z_3) \cdot (z_2, z_4) = (z_1, z_4) \cdot (z_2, z_3) = (z_1, z_2) \cdot (z_3, z_4)$$

It is not difficult to see that *this condition completely characterizes an equianharmonic quadrangle*: in fact, it follows from this condition that $|\lambda| = 1$ and $|1 - \lambda| = 1$, where

$$\lambda = W(z_1, z_2, z_3, z_4),$$

and this is possible only if $\lambda = (1 + \sqrt{3}i)/2$ or $\lambda = (1 - \sqrt{3}i)/2$. In particular, a rhombus is an equianharmonic quadrangle, if the product of the diagonals is equal to the square of a side (a rhombus with side 1 and diagonals $(\sqrt{3} + 1)/\sqrt{2}$ and $(\sqrt{3} - 1)/\sqrt{2}$).

For any triangle $\overline{z_1z_2z_3}$ it is possible to find two points z_4 and z_4' such that the quadrangles $\overline{z_1z_3z_2z_4}$ and $\overline{z_1z_3z_2z_4'}$ are equianharmonic; these are *the points of intersection of the circle of Apollonius S_3 of the points z_1 and z_2, which passes through z_3, and the circle of Apollonius S_2 of the points z_1 and z_3, which*

passes through z_2 (hence it follows that *through these points there passes the circle of Apollonius* S_1 *of the points* z_2 *and* z_3, *which passes through* z_1). In fact, from the equations

$$\frac{(w, z_1)}{(w, z_2)} = \frac{(z_3, z_1)}{(z_3, z_2)} \quad \text{and} \quad \frac{(w, z_1)}{(w, z_3)} = \frac{(z_2, z_1)}{(z_2, z_3)}$$

it follows that the equation

$$\frac{(w, z_2)}{(w, z_3)} = \frac{(z_1, z_2)}{(z_1, z_3)}$$

also is satisfied, and that

$$(z_1, z_2) \cdot (w, z_3) = (z_1, z_3) \cdot (w, z_2) = (z_2, z_3) \cdot (w, z_1)$$

The points z_4 and z_4' may be defined as the points in the plane of the triangle $\overline{z_1z_2z_3}$ whose distances from any two vertices are in the same ratio as the lengths of the corresponding sides (the distances of the third vertex from these two vertices); these are sometimes called the **isodynamic centers** of the triangle $\overline{z_1z_2z_3}$. Each triangle has two isodynamic centers, z_4 and z_4'; only for an equilateral triangle $\overline{z_1z_2z_3}$, for which both circles of Apollonius S_3 and S_2 (and also the circle S_1) reduce to a line, is there a *unique* isodynamic center w_4, which coincides with the center of the triangle. In this case, it may be noted, the role of the center w_4' is played by the point at infinity ∞, since

$$W(z_1, z_2, z_3, \infty) = V(z_1, z_2, z_3) = \frac{1 + \sqrt{3}i}{2} \text{ or } \frac{1 - \sqrt{3}i}{2}$$

The six values of the ratio of three points of the plane can be obtained more geometrically by using a diagram rather than calculation. We note, first, that the ratio $V(z_1, z_2, z_3) = (z_1 - z_3)/(z_2 - z_3)$ characterizes a triad of points z_1, z_2, and z_3 "to within a similarity"; this phrase means that two triads of points, z_1, z_2, z_3 and z_1', z_2', z_3', have the same ratio, $V(z_1, z_2, z_3) = V(z_1', z_2', z_3')$, if and only if the triangles $\overline{z_1z_2z_3}$ and $\overline{z_1'z_2'z_3'}$ formed by these points are similar. (The reason is that under this similarity the point z_1' must correspond to z_1, the point z_2'

to z_2, and the point z'_3 to z_3; we note that the "triangles" $\overline{z_1z_2z_3}$ and $\overline{z'_1z'_2z'_3}$ can be degenerate, since the points z_1, z_2, and z_3 can lie on one line.) This statement follows immediately from the fact that the modulus $|V|$ of the complex number $V(z_1, z_2, z_3) = (z_1 - z_3)/(z_2 - z_3)$ is equal to the ratio of the lengths of the sides $\overline{z_1z_3}$ and $\overline{z_2z_3}$ of the triangle $\overline{z_1z_2z_3}$, and its argument, arg V, is equal to the angle $\angle\{[z_3z_2], [z_3z_1]\}$ of the triangle.

But it is clear that

$$V(z, 1, 0) = \frac{z - 0}{1 - 0} = z$$

Therefore, if $V(z_1, z_2, z_3) = \lambda = V(\lambda, 1, 0)$, then the triangle $\overline{z_1z_2z_3}$ is similar to the triangle $\overline{\lambda 1 O}$. Hence it follows that to find all the values of the ratio of the three points z_1, z_2, and z_3 taken in any order it is necessary to construct on the segment $\overline{O1}$ all possible triangles similar to the given triangle $\overline{z_1z_2z_3}$. This can be done in six different ways: $\overline{\lambda 1 O} \sim \overline{z_1z_2z_3}$, $\overline{\lambda_1 1 O} \sim \overline{z_2z_1z_3}$, $\overline{\lambda_2 1 O} \sim \overline{z_1z_3z_2}$, $\overline{\lambda_3 1 O} \sim \overline{z_3z_1z_2}$, $\overline{\lambda_4 1 O} \sim \overline{z_2z_3z_1}$, and $\overline{\lambda_5 1 O} \sim \overline{z_3z_2z_1}$, where the correspondence between the vertices is indicated by the order in which they are written (Figure 26).

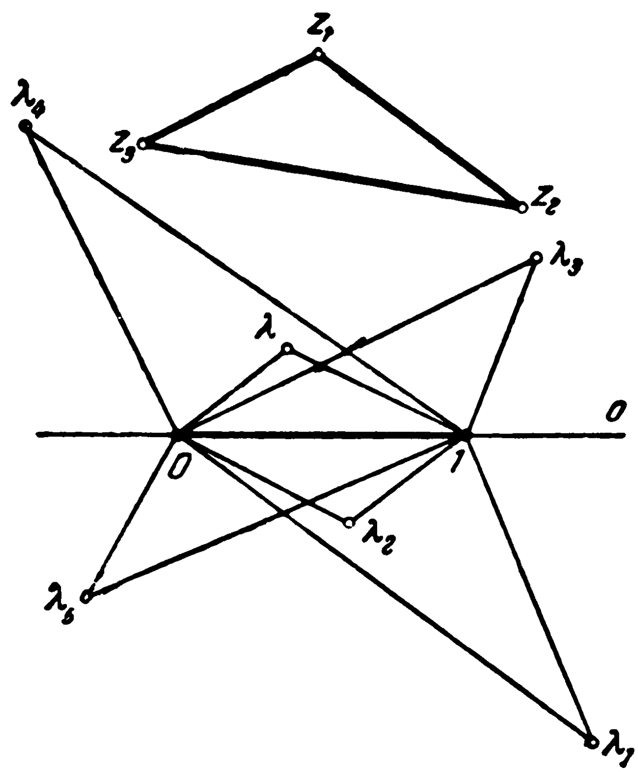


FIG. 26

From the similarity of the triangles $\overline{1\lambda_1 O}$, $\overline{\lambda_2 O 1}$, $\overline{1 O \lambda_3}$, $\overline{O\lambda_4 1}$, and $\overline{O 1 \lambda_5}$ to the triangle $\overline{\lambda 1 O}$ it can be deduced that

$$\frac{\lambda_1}{1} = \frac{1}{\lambda}, \text{ i.e. } \lambda_1 = \frac{1}{\lambda}, \qquad \frac{\lambda_2}{1} = \frac{1-\lambda}{1} = 1 - \lambda,$$

$$\frac{\lambda_3}{1} = \frac{0-1}{\lambda - 1} = \frac{1}{1-\lambda}, \qquad \frac{\lambda_4}{1} = \frac{1-\lambda}{0-\lambda} = \frac{\lambda - 1}{\lambda},$$

$$\frac{\lambda_5}{1} = \frac{0-\lambda}{1-\lambda} = \frac{\lambda}{\lambda - 1}$$

To explain under what circumstances the ratio of three points takes less than six different values it is sufficient to determine under what circumstances any two of the six triangles represented in Figure 26 coincide. One such case is quite obvious. It occurs when the triangle $\overline{O1\lambda}$ is equilateral, and instead of six triangles we have only two; the points λ, λ_3, and λ_4 and also the points λ_1, λ_2, and λ_5 coincide (cf. Figure 24b). In this case

$$\lambda = \lambda_3 = \lambda_4 = 1(\cos 60^\circ + i \sin 60^\circ) = \frac{1 + \sqrt{3}i}{2}$$

$$\lambda_1 = \lambda_2 = \lambda_5 = 1[\cos(-60^\circ) + i \sin(-60^\circ)] = \frac{1 - \sqrt{3}i}{2}$$

A second case, which also can be seen with no difficulty, occurs when the vertices λ and λ_2 of the parallelogram $\overline{O\lambda 1 \lambda_2}$ coincide. It is clear that this is possible only when $\lambda = \lambda_2 = \frac{1}{2}$ and the point λ is the midpoint of the segment $\overline{O1}$, so that $\lambda_1 = \lambda_3 = 2$ and $\lambda_4 = \lambda_5 = -1$, and the ratio of the three points z_1, z_2, and z_3 takes only the three values -1, 2, and $\frac{1}{2}$; cf. Figure 24a. It is easy to see from Figure 26 that these two cases, in which the number of triangles is reduced, are *the only ones*.

Similarly, in the case of four points, z_1, z_2, z_3, and z_4, instead of the six values λ, λ_1, λ_2, λ_3, λ_4, and λ_5 of the cross-ratio W of these points we may consider a triangle $\overline{O1\lambda}$ defined "to within a similarity"; that is, two such triangles which are similar to each other (this time the order of their vertices not being taken into account) are not distinct and are identified with each other. The latter condition associates with each triangle $\overline{O1\lambda}$ five more "identical" triangles $\overline{O1\lambda_1}$, $\overline{O1\lambda_2}$, $\overline{O1\lambda_3}$, $\overline{O1\lambda_4}$, and $\overline{O1\lambda_5}$; see Figure 26. The triangle $\overline{O1\lambda}$ is closely connected with the quadrangle $\overline{z_1 z_2 z_3 z_4}$; we call it the **associated triangle** of the quadrangle $\overline{z_1 z_2 z_3 z_4}$. The connection between the quadrangle $\overline{z_1 z_2 z_3 z_4}$ and the triangle $\overline{O1\lambda}$ enables us to deduce many properties of the quadrangle

from (simpler) properties of its associated triangle. Thus, for example, to a degenerate triangle $\overline{O1\lambda}$ (that is, one whose vertices lie on one line) corresponds a quadrangle $\overline{z_1z_2z_3z_4}$ which may be inscribed in a circle (more precisely, one whose vertices lie on one circle or on one line); only for such quadrangles $\overline{z_1z_2z_3z_4}$ is the cross-ratio $W(z_1, z_2, z_3, z_4)$ real. Later we shall have occasion to dwell at greater length on the properties of the associated triangle of a given quadrangle and to indicate its construction purely geometrically (Section 14). Here we merely mention the fact that the associated triangle of a harmonic quadrangle degenerates into a segment and its midpoint, and the associated triangle of an equianharmonic quadrangle is equilateral.

§9. Dual Numbers as Oriented Lines of a Plane

Later we shall need to deal exclusively with oriented lines; we shall often drop the adjective "oriented." The angle between the lines a and b will be called the **oriented angle** $\angle\{a, b\}$ between the oriented lines (see p. 31); the distance between two points A and B of a line l will be called the **oriented length** of the segment $\overline{AB}$, denoted by $\{A, B\}$ and meaning the ordinary distance, taken with a plus or a minus sign, depending on whether the direction from A to B coincides with the positive or the negative direction of the line l; the distance from a point M to an oriented line l will be called the **oriented distance** $\{M, l\}$ from M to l, which is the distance taken with a plus or a minus sign, depending on whether M lies to the left or to the right of the oriented line l. We shall call two oriented lines *parallel* only if they are parallel in the ordinary sense and their directions coincide (Figure 27a); we shall sometimes call parallel lines with opposite directions *antiparallel* (Figure 27b). The distance from a line a to a line b which does not intersect it will be called the **oriented distance** $\{a, b\}$ from a to b, that is, the

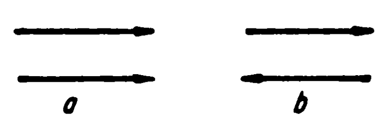


FIG. 27

oriented distance from an arbitrary point of the line a to the line b; obviously, $\{a, b\} = -\{b, a\}$ if a and b are parallel, and $\{a, b\} = \{b, a\}$ if a and b are antiparallel.

We now recall that the polar coordinates of points of a plane are defined by taking some point O (the *pole* of the coordinate system) and an (oriented) line o (the *polar axis*) passing through O; the coordinates of a point M are the distance $r = OM$ of this point from the pole and the angle $\varphi = \angle\{o, m\}$ which the (oriented) line m joining O and M makes with o (cf. Figure 1). Similarly we may define the **polar coordinates** of (oriented) lines of the plane, for which it is necessary to take some (oriented) line o (the polar axis) and a point O (the pole) lying on o; the coordinates of a line l are the angle $\theta = \angle\{o, l\}$ which l makes with the polar axis o and the (oriented) distance $s = \{O, L\}$ from O to the point of intersection L of l and o; see Figure 28a. Obviously,

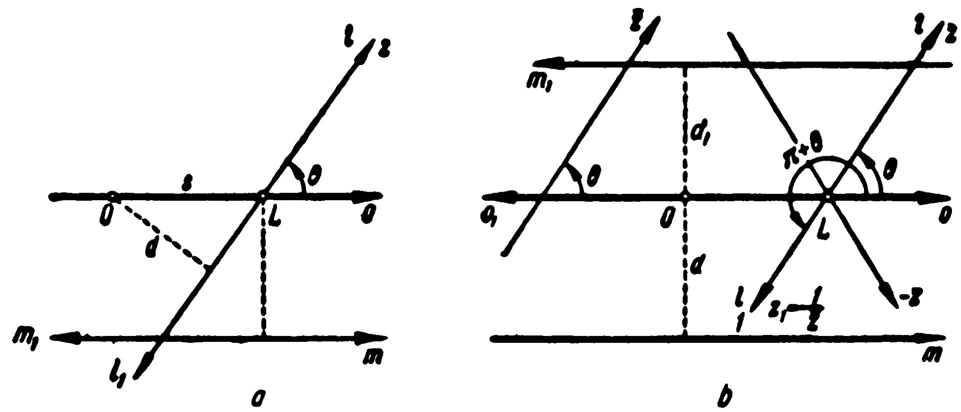


FIG. 28

the coordinate s of an oriented line l can have any value between $+\infty$ and $-\infty$; the coordinate θ can have any value between 0 and 2π. It is natural to say that $\theta = 0$ for lines parallel to the polar axis o and $\theta = \pi$ for lines antiparallel to o; if a line does not intersect the axis o, it does not have a coordinate s (we may say that in such case $s = \pm\infty$).

Let us now agree to associate with an (oriented) line l with polar coordinates θ and s the *dual number*

$$z = \tan\frac{\theta}{2}\cdot(1 + \varepsilon s) = u + \varepsilon v, \qquad u = \tan\frac{\theta}{2}, \qquad v = \tan\frac{\theta}{2}\cdot s \tag{29}$$

See Figure 28. Here lines parallel to o, for which $\theta = 0$, naturally correspond to numbers of zero modulus, that is, to divisors of zero εv (see Section 4, Equation 31). In order to establish an exact correspondence between lines parallel to o and divisors of zero we note that the distance $d = \{O, l\}$ of a line l not parallel to o from the pole O is equal to

$$d = s \sin \theta = s \frac{2 \tan \theta/2}{1 + \tan^2 \theta/2} = \frac{2s \tan \theta/2}{1 + \tan^2 \theta/2} = \frac{2v}{1 + | z |^2} \tag{30}$$

See Figure 28a. If we want Equation 30 to remain valid for a line m *parallel to* o and separated from o by a distance $\{o, m\} = d$, then this line will need to correspond to the number $z = (d/2)\varepsilon$, that is, to $z = u + \varepsilon v$, where $u = 0$ and $2v/(1 + | z |^2) = 2v = d$.

Two lines l and l_1, which intersect o and *differ only in direction*, and so have polar coordinates (θ, s) and $(\pi + \theta, s)$, correspond to dual numbers

$$z = \tan \frac{\theta}{2} (1 + \varepsilon s)$$

and

$$\begin{aligned} z &= \tan \frac{\pi + \theta}{2} (1 + \varepsilon s) = -\cot \frac{\theta}{2} (1 + \varepsilon s) \\ &= -\frac{1}{\tan \theta/2 \cdot (1 - \varepsilon s)} = -\frac{1}{\bar{z}} \end{aligned}$$

Assuming that this relation remains valid for lines which do not intersect o, we agree to assign to a line m_1 *antiparallel to* o and separated from o by a distance $\{o, m_1\} = d_1$ the number

$$z = -\frac{1}{\overline{-\frac{1}{2} d_1 \varepsilon}} = -\frac{2}{d_1} \omega$$

We may note that, if the distance $\{o, m\}$ from o to a line m, which is parallel to o and coincides in position with the line m_1, is equal to d, then $d = -d_1$. Finally, with the line o_1, the *antiaxis*, which differs only in direction from the polar axis o, we associate the number

$$\frac{1}{0} = \infty$$

In this way we establish a complete correspondence between oriented lines of a plane and dual numbers, including both numbers of the form $w\omega$, where $w \neq 0$ is real, and the number ∞.

It is obvious that to real numbers $z = u = \tan \theta/2 \cdot (1 + \varepsilon \cdot 0)$ correspond lines passing through the pole O. To numbers of modulus 1 correspond lines perpendicular to o. More precisely, to numbers of modulus 1 correspond lines l such that $\{o, l\} = \pi/2$; more generally, to numbers of constant modulus u correspond lines l which make with o a constant angle $\angle\{o, l\} = 2 \tan^{-1} u$. To a "purely imaginary" number $v\varepsilon$ (a number of zero modulus) and "numbers of infinite modulus" $w\omega$ correspond lines parallel and antiparallel to the axis o. To conjugate numbers $z = \tan \theta/2 \cdot (1 + \varepsilon s)$ and $\bar{z} = \tan \theta/2 \cdot (1 - \varepsilon s)$ correspond lines symmetrical about the pole O. To opposite numbers $z = \tan \theta/2 \cdot (1 + \varepsilon s)$ and $-z = -\tan \theta/2 \cdot (1 + \varepsilon s) = \tan (2\pi - \theta)/2 \cdot (1 - \varepsilon s)$ correspond lines symmetrical about the polar axis o, that is, lines which meet o in the same point L and make equal angles with o: $\angle\{o, z\} = \angle\{-z, o\}$; see Figure 28b. To numbers z and $-1/\bar{z}$ correspond lines which differ only in direction. Thus, the equations

$$z' = \bar{z}, \qquad z' = -z, \qquad z' = -\frac{1}{\bar{z}} \tag{31a,b,c}$$

may be regarded as defining transformations in the set of oriented lines of the plane: *symmetry about the point O*, *symmetry about the line o*, and *reorientation* (change of direction of all lines of the plane), respectively.

We now explain how to write down, with the help of dual numbers, arbitrary **motions** (among which we shall include reorientation, which also leaves distances between points of the plane unaltered). First, it is clear that the *translation along o* through a distance t takes the line which corresponds to the dual number

$$z = \tan \frac{\theta}{2} \cdot (1 + \varepsilon s)$$

into the line which corresponds to the number

$$z' = \tan \frac{\theta'}{2} \cdot (1 + \varepsilon s') = \tan \frac{\theta}{2} \cdot [1 + \varepsilon(s + t)]$$

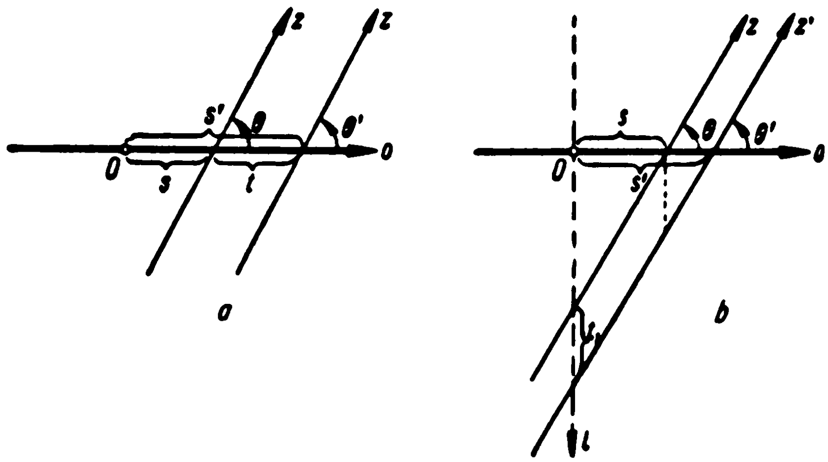


FIG. 29

See Figure 29a. Subsequently in such cases we shall say briefly, "takes the line $z = \tan \theta/2 \cdot (1 + \varepsilon s)$ into the line $z' = \tan \theta/2 \cdot [1 + \varepsilon(s + t)]$." Now, it follows that this translation may be written as

$$z' = pz, \qquad p = 1 + \varepsilon t, \quad |p| = 1 \tag{32}$$

since $[1 \cdot (1 + \varepsilon t)] \cdot [\tan \theta/2 \cdot (1 + \varepsilon s)] = \tan \theta/2 \cdot [1 + \varepsilon(s + t)]$. Further, *the translation* through a distance t_1 *in the direction perpendicular to o* (in the direction of a line l such that $\angle\{l, o\} = \pi/2$) takes the line

$$z = \tan \frac{\theta}{2} \cdot (1 + \varepsilon s)$$

into the line

$$z' = \tan \frac{\theta'}{2} \cdot (1 + \varepsilon s') = \tan \frac{\theta}{2} \cdot [1 + \varepsilon(s + t_1 \cot \theta)]$$

See Figure 29b. But

$$\begin{aligned} z' &= \tan \frac{\theta}{2} \cdot [1 + \varepsilon(s + t_1 \cot \theta)] \\ &= \tan \frac{\theta}{2} (1 + \varepsilon s) + \varepsilon \tan \frac{\theta}{2} \cdot \frac{t_1(1 - \tan^2 \theta/2)}{2 \tan \theta/2} \\ &= \tan \frac{\theta}{2} (1 + \varepsilon s) + \varepsilon \frac{t_1}{2} - \left(\varepsilon \frac{t_1}{2} \tan^2 \frac{\theta}{2}\right) = z + \varepsilon \frac{t_1}{2} - \varepsilon \frac{t_1}{2} z^2 \end{aligned}$$

The last formula may be written in a more elegant form. We note that

$$z + \varepsilon\frac{t_1}{2} - \varepsilon\frac{t_1}{2}z^2 = \left(z + \varepsilon\frac{t_1}{2}\right)\left(1 - \varepsilon\frac{t_1}{2}z\right)$$
$$= \left(z + \varepsilon\frac{t_1}{2}\right)\Big/\left(\varepsilon\frac{t_1}{2}z + 1\right)$$

Thus, the given translation may be denoted by the formula

$$z' = \frac{z + q}{qz + 1}, \qquad q = \varepsilon\frac{t_1}{2}, \quad |\, q\,| = 0 \tag{32a}$$

Hence it follows that an *arbitrary translation* (that is, a translation through a distance t in the direction of o and through a distance t_1 in the direction of l perpendicular to o) may be denoted by the formula

$$z' = \frac{(pz) + q}{q(pz) + 1}, \qquad p = 1 + \varepsilon t, \quad q = \varepsilon\frac{t_1}{2}$$

or, if we introduce the notation $p = p_1^2$, in which $p_1 = 1 + \varepsilon(t/2)$, and use the fact that $q = \varepsilon(t_1/2) = \varepsilon(t_1/2)[1 + \varepsilon(t/2)] = qp_1$ and $\bar{p}_1 = 1 - \varepsilon(t/z) = 1/p_1$ and $\bar{q} = -q$, it may be denoted by the formula

$$z' = \frac{p_1^2 z + qp_1}{qp_1^2 z + p_1\bar{p}_1} = \frac{p_1 z + q}{qp_1 z + \bar{p}_1} = \frac{p_1 z + q}{-\bar{q}z + \bar{p}_1} \tag{33}$$

where $p_1 = 1 + \varepsilon(t/2)$, $q = \varepsilon(t_1/2)$, $|\, p_1\,| = 1$, and $|\, q\,| = 0$.

We now turn to *rotations* of the plane. It is obvious that the rotation about O through an angle α takes the line $z = \tan\theta/2 \cdot (1 + \varepsilon s)$ into the line $z' = \tan\theta'/2 \cdot (1 + \varepsilon s')$, where $\theta' = \theta + \alpha$; see Figure 30. Thus,

$$|\, z'\,| = \tan\frac{\theta + \alpha}{2} = \frac{\tan\theta/2 + \tan\alpha/2}{1 - (\tan\theta/2 \cdot \tan\alpha/2)}$$
$$= \frac{|\, z\,| + \tan\alpha/2}{-(\tan\alpha/2 \cdot |\, z\,|) + 1} = \left|\frac{z + \tan\alpha/2}{-(\tan\alpha/2 \cdot z) + 1}\right| \tag{34}$$

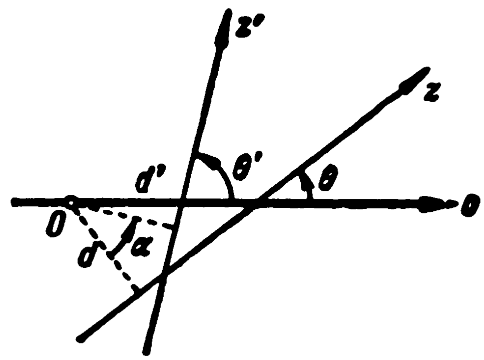


Fig. 30

(Here we use the fact that if z_1 and z_2 are dual numbers, then $|z_1 \pm z_2| = |z_1| \pm |z_2|$, $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$, and $|z_1/z_2| = |z_1|/|z_2|$.) Further, if d and d' are the distances of the lines z and z' from the pole O, then (cf. Equation 30):

$$s \sin \theta = d = d' = s' \sin \theta'$$

Therefore

$$\arg z' = s' = s \frac{\sin \theta}{\sin \theta'} = s \frac{\sin \theta}{\sin(\theta + \alpha)}$$

On the other hand, since $\arg(u + \varepsilon v) = v/u$,

$$\begin{aligned}
\arg \frac{z + \tan \alpha/2}{-(\tan \alpha/2 \cdot z) + 1} &= \arg \left[z + \tan \frac{\alpha}{2}\right] \\
&\quad - \arg \left[-\left(\tan \frac{\alpha}{2} \cdot z\right) + 1\right] \\
&= \arg \left[\tan \frac{\theta}{2} (1 + \varepsilon s) + \tan \frac{\alpha}{2}\right] \\
&\quad - \arg \left[-\left(\tan \frac{\alpha}{2} \cdot \tan \frac{\theta}{2}\right) (1 + \varepsilon s) + 1\right] \\
&= \frac{s \tan \theta/2}{\tan \theta/2 + \tan \alpha/2} \\
&\quad - \frac{-s(\tan \alpha/2 \cdot \tan \theta/2)}{1 - (\tan \alpha/2 \cdot \tan \theta/2)} \qquad (34a) \\
&= \frac{\tan \theta/2 \cdot (\tan^2 \alpha/2 + 1)}{(\tan \theta/2 + \tan \alpha/2)[1 - (\tan \alpha/2 \cdot \tan \theta/2)]} s
\end{aligned}$$

$$= \frac{(\tan\theta/2 \cdot \sec^2\alpha/2)(\cos\theta/2 \cdot \cos\alpha/2)^2}{\sin(\alpha+\theta)/2 \cdot \cos(\alpha+\theta)/2} s$$

$$= \frac{\sin\theta/2 \cdot \cos\theta/2}{\sin(\alpha+\theta)/2 \cdot \cos(\alpha+\theta)/2} s$$

$$= \frac{\sin\theta}{\sin(\alpha+\theta)} s = \arg z'$$

From Equations 34 and 34a it follows that our rotation is denoted by the formula

$$z' = \frac{z + q_1}{-\bar{q}_1 z + 1} \tag{35}$$

where[28]

$$q_1 = \bar{q}_1 = \tan\alpha/2 \quad \text{and} \quad \arg q_1 = 0$$

Finally, the **most general motion** (cf. Equation 4) consists of a rotation (Equation 35) about O through some angle α together with a translation (Equation 33):

$$z' = \frac{p_1(z+q_1)/(-\bar{q}_1 z + 1) + q}{-\bar{q}(z+q_1)/(-\bar{q}_1 z + 1) + p_1}$$

$$= \frac{(p_1 - q\bar{q}_1)z + (p_1 q_1 + q)}{-(\bar{p}_1\bar{q}_1 + \bar{q})z + (\bar{p}_1 - \bar{q}q_1)}$$

This transformation may be written in the form

$$z' = \frac{Pz + Q}{-\bar{Q}z + \bar{P}} \tag{36a}$$

where

$$P = p_1 - q\bar{q} \quad \text{and} \quad Q = p_1 q_1 + q$$

It is possible that the original motion consists of the symmetry

[28] In this formula the number $q_1 = \tan\alpha/2$ can also take the value ∞, which corresponds to a rotation through an angle of 180°. If $q_1 = \infty$, then we have $z' = (z + \infty)/(-\infty \cdot z + 1) = -1/z$; (see p. 16). This transformation obviously reduces to the symmetry about the point O, Equation 31a, followed by the reorientation, Equation 31c. (We note that the symmetry about the point O, defined above, *does not* coincide with the rotation through 180° about O.)

(Equation 31b) about the line o together with a rotation about O and a translation (Equation 36a):

$$z' = \frac{-Pz + Q}{\bar{Q}z + \bar{P}} \tag{36b}$$

Finally, the motion may consist of the reorientation (Equation 31c) together with one of the transformations (Equation 36a or 36b):

$$z' = \frac{-P(1/\bar{z}) + Q}{\bar{Q}(1/\bar{z}) + \bar{P}} = \frac{P_1\bar{z} + Q_1}{-\bar{Q}_1\bar{z} + \bar{P}_1} \tag{36c}$$

where $P_1 = Q$ and $Q_1 = -P$, or

$$z' = \frac{P(1/\bar{z}) + Q}{-\bar{Q}(1/\bar{z}) + \bar{P}} = \frac{-P_1\bar{z} + Q_1}{\bar{Q}_1\bar{z} + \bar{P}_1} \tag{36d}$$

where $P_1 = -Q$ and $Q_1 = P$.

It is obvious that the (oriented) **angle** $\delta = \angle\{z_1, z_2\}$ *between the lines* $z_1 = \tan\theta_1/2 \cdot (1 + \varepsilon s_1)$ *and* $z_2 = \tan\theta_2/2 \cdot (1 + \varepsilon s_2)$ *is equal to* $\theta_2 - \theta_1$; see Figure 31a. This may be written as

$$\tan\frac{\delta}{2} = \frac{\tan\theta_2/2 - \tan\theta_1/2}{1 + (\tan\theta_2/2 \cdot \tan\theta_1/2)}$$

$$= \frac{|z_2| - |z_1|}{1 + (|z_2| \cdot |z_1|)} = \left|\frac{z_2 - z_1}{1 + z_1\bar{z}_2}\right|$$

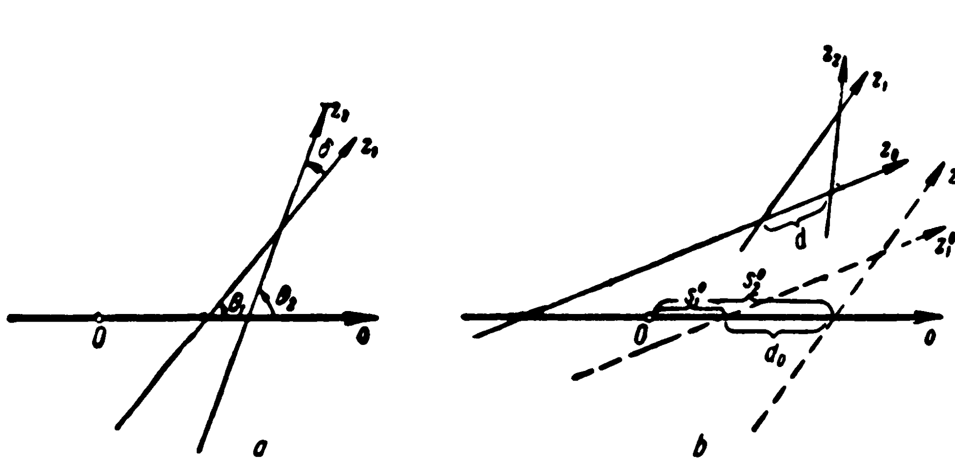


FIG. 31

The result thus obtained can be put into the following form:

$$\tan^2 \frac{\delta}{2} = \frac{(z_2 - z_1)(\bar{z}_2 - \bar{z}_1)}{(1 + z_1\bar{z}_2)(1 + \bar{z}_1 z_2)} \tag{37}$$

We now find the (oriented) **distance** $d = \{[z_1 z_0], [z_2 z_0]\}$ *between the points of intersection* $[z_1 z_0]$ *and* $[z_2 z_0]$ *of a definite line* z_0 *with two other lines* z_1 *and* z_2 ; see Figure 31b. It is obvious that the distance d_0 between the points of intersection of the line o and the lines $z_1 = \tan \theta_1^0/2 \cdot (1 + \varepsilon s_1^0)$ and $z_2^0 = \tan \theta_2^0/2 \cdot (1 + \varepsilon s_2^0)$ is equal to

$$d_0 = \arg z_2^0/z_1^0 , \qquad (= \arg z_2^0 - \arg z_1^0 = s_2^0 - s_1^0)$$

An example of a motion which takes a given line z_0 into the line o is given by the formula

$$z' = \frac{z - z_0}{\bar{z}_0 z + 1}$$

This motion takes the lines z_1 and z_2 into the lines $z_1^0 = (z_1 - z_0)/(\bar{z}_0 z_1 + 1)$ and $z_2^0 = (z_2 - z_0)/(\bar{z}_0 z_2 + 1)$. Hence we obtain[29]

$$d = \arg \frac{(z_2 - z_0) : (\bar{z}_0 z_2 + 1)}{(z_1 - z_0) : (\bar{z}_0 z_1 + 1)} = \arg \left(\frac{z_2 - z_0}{z_1 - z_0} : \frac{\bar{z}_0 z_2 + 1}{\bar{z}_0 z_1 + 1} \right)$$

$$= \arg \frac{z_2 - z_0}{z_1 - z_0} - \arg \frac{\bar{z}_0 z_2 + 1}{\bar{z}_0 z_1 + 1} \tag{38}$$

The condition that the three lines z_0 , z_1 , *and* z_2 *should meet in one point is that* the distance between the points of intersection of z_1 and z_0 with z_2 should be equal to zero, that is, by virtue of Equation 38, *the ratio* $[(z_0 - z_2)/(z_1 - z_2)]/[(\bar{z}_2 z_0 + 1)/(\bar{z}_2 z_1 + 1)]$ *should be real.* This condition may be written as

$$\frac{(z_0 - z_2)(\bar{z}_2 z_1 + 1)}{(z_1 - z_2)(\bar{z}_2 z_0 + 1)} = \frac{(\bar{z}_0 - \bar{z}_2)(z_2 \bar{z}_1 + 1)}{(\bar{z}_1 - \bar{z}_2)(z_2 \bar{z}_0 + 1)} \tag{39}$$

[29] If one of the lines z_1 and z_2 (or even both these lines) does not intersect z_0 (or coincides in position with z_0), then one of the numbers $(z_2 - z_0)/(\bar{z}_0 z_2 + 1)$ and $(z_1 - z_0)/(\bar{z}_0 z_1 + 1)$ (and possibly both numbers) does not reduce to the form in Equation 32 of Section 4 (that is, it is a divisor of zero or a number reciprocal to it, and so does not have an argument).

Consequently, the *equation of a point*, which is the condition that a line z must satisfy if it is to pass through the point $[z_1 z_2]$, has the form

$$\frac{(z - z_2)(\bar{z}_2 z_1 + 1)}{(z_1 - z_2)(\bar{z}_2 z + 1)} = \frac{(\bar{z} - \bar{z}_2)(z_2 \bar{z}_1 + 1)}{(\bar{z}_1 - \bar{z}_2)(z_2 \bar{z} + 1)}$$

or[30]

$$Az\bar{z} + Bz - \overline{B}\bar{z} - A = 0, \qquad A \text{ purely imaginary} \quad (40)$$

Here $A = z_2(\bar{z}_1 - \bar{z}_2)(\bar{z}_2 z_1 + 1) - \bar{z}_2(z_1 - z_2)(z_2 \bar{z}_1 + 1) = (z_2 \bar{z}_1 - z_1 \bar{z}_2)(1 + z_2 \bar{z}_2)$, and $B = (\bar{z}_1 - \bar{z}_2)(\bar{z}_2 z_1 + 1) + \bar{z}_2^2(z_1 - z_2)(z_2 \bar{z}_1 + 1)$. Conversely, it is not difficult to see that *every equation of the form of Equation* 40 *represents a point.*[30a]

We now find the condition that four oriented lines z_0, z_1, z_2, and z_3 should touch one oriented circle. By **oriented circle** we understand the set ("locus") of all oriented lines l, whose oriented distance $\{O, l\}$ from a given point O (the *center* of the circle) has a fixed value r. The number r is called the *radius* of the circle; thus, *the radius of an oriented circle may be positive or negative* (if $r = 0$, the oriented circle degenerates to a point, which is thus a special case of a circle). In a diagram an oriented circle is represented as an ordinary circle provided with an arrow, which denotes a definite direction of rotation, the direction at each point coinciding with the direction of the tangent to the circle at that point; see Figure 32. From the definition of the oriented distance $\{O, l\}$ from a point O to a line l (p. 80) it follows

---

[30] Equation 40 follows also from the fact that by the translation, Equation 33, any point M can be taken into the pole O, whose equation has the form $z - \bar{z} = 0$. Therefore the equation of M may be written as $(p_1 z + q)/(-\bar{q}z + \bar{p}_1) - (\bar{p}_1 \bar{z} + \bar{q})/(-q\bar{z} + p_1) = 0$, that is, in the form of Equation 40, where $A = \bar{p}_1 \bar{q} - p_1 q = -\varepsilon t_1$ and $B = p^2 + \bar{q}^2 = p^2 = 1 + \varepsilon t$; here t and t_1 are the components of translation, in the direction of the axis o and the direction perpendicular to o, which take M into O ($-t$ and $-t_1$ are the rectangular coordinates of the point M).

[30a] In order that the last statement should be completely accurate, we must agree to include among the "points" pencils of lines perpendicular to the axis o (we can regard these lines as passing through a fictitious point at infinity; the equation of the pencil can be written in the form $Az\bar{z} - A = 0$, where A is any purely imaginary number).

that the radius of an oriented circle is positive if the direction of rotation is counterclockwise and is negative otherwise.

It can be shown that *four* (oriented) *lines* z_0, z_1, z_2, *and* z_3 *touch one* (oriented) *circle or pass through one point if and only if*[31]

$$\{[z_0z_2], [z_1z_2]\} + \{[z_1z_3], [z_0z_3]\} = \{[z_3z_0], [z_2z_0]\} + \{[z_2z_1], [z_3z_1]\} \tag{41}$$

In order to verify this, let us consider Figure 33, which shows four (oriented) tangents z_0, z_1, z_2, and z_3 to an (oriented) circle S (if the lines z_0, z_1, z_2, and z_3 pass through one point, then the condition, Equation 41, is clearly satisfied), touching S at points M, N, P, and Q, respectively; the points $[z_0z_2]$, $[z_1z_2]$,

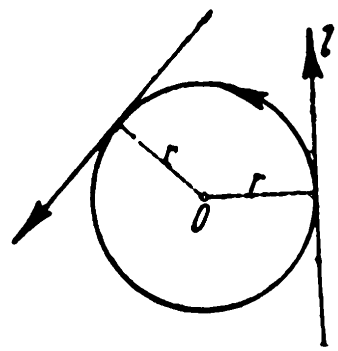


FIG. 32

$[z_1z_3]$, and $[z_0z_3]$ are denoted, for brevity, by A, B, C, and D. Here, obviously (taking into account the fact that all the distances which occur below are oriented), we have[32]

$$\{A, B\} + \{C, D\} = \{A, P\} + \{P, B\} + \{C, Q\} + \{Q, D\}$$
$$\{D, A\} + \{B, C\} = \{D, M\} + \{M, A\} + \{B, N\} + \{N, C\}$$

But since, by virtue of a well-known property of tangents to a circle,

$$\{A, P\} = \{M, A\}, \qquad \{P, B\} = \{B, N\}$$
$$\{C, Q\} = \{N, C\}, \qquad \{Q, D\} = \{D, M\}$$

[31] Note the condition that the four *points* z_0, z_1, z_2, and z_3 should lie on one circle. This condition may be written as $\angle\{[z_0z_2], [z_1z_2]\} + \angle\{[z_1z_3], [z_0z_3]\} = \angle\{[z_3z_0], [z_2z_0]\} + \angle\{[z_2z_1], [z_3z_1]\}$; see the beginning of Section 7, particularly Figure 6.

[32] Here we make use of the fact that if X, Y, and Z are three points of an oriented line, then *in all cases* $\{X, Y\} + \{Y, Z\} = \{X, Z\}$.

then the condition, Equation 41, is satisfied in all cases:[33]

$$\{A, B\} + \{C, D\} = \{D, A\} + \{B, C\}$$

It is not difficult to see that, conversely, *if the condition* (Equation 41) *is satisfied, then the four* (not all parallel) *lines* z_0, z_1, z_2, *and* z_3 *touch one oriented circle or pass through one point.*[34]

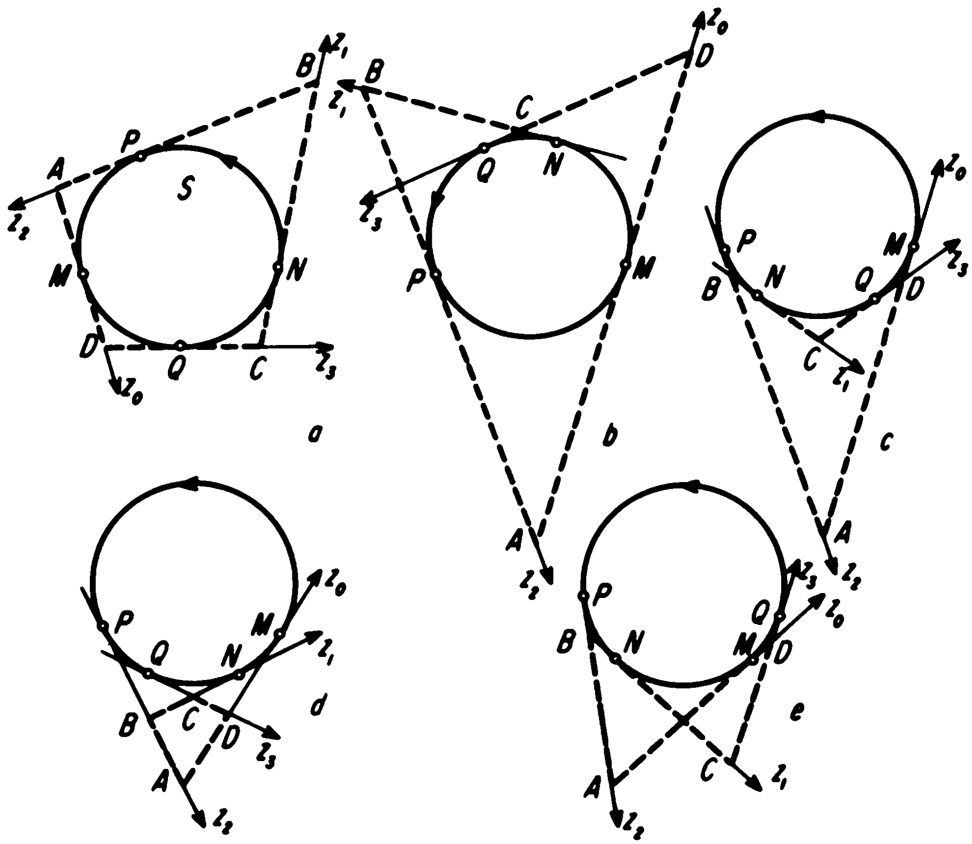


FIG. 33

[33] This is the precise formulation of a well-known theorem of geometry on the equality of the sums of opposite sides of an inscribed quadrangle. If the distances are not oriented, the formulation of the theorem is considerably complicated; in such case we have either the equality $AB + CD = AD + BC$ (Figure 33a, b) or the equality $AB - CD = AD - BC$ (Figure 33c-e).

[34] See, for example, A. Clement Jones, *A Geometry for Schools*, p. 154 (Arnold), 1920, or C. V. Durell and A. Robson, *Advanced Trigonometry*, p. 28 (Bell), 1930.

Using Equation 38, we may rewrite the condition (Equation 41) in the form

$$\arg\frac{z_1 - z_2}{z_0 - z_2} - \arg\frac{\bar{z}_2 z_1 + 1}{\bar{z}_2 z_0 + 1} + \arg\frac{z_0 - z_3}{z_1 - z_3} - \arg\frac{\bar{z}_3 z_0 + 1}{\bar{z}_3 z_1 + 1}$$
$$= \arg\frac{z_2 - z_0}{z_3 - z_0} - \arg\frac{\bar{z}_0 z_2 + 1}{\bar{z}_0 z_3 + 1} + \arg\frac{z_3 - z_1}{z_2 - z_1} - \arg\frac{\bar{z}_1 z_3 + 1}{\bar{z}_1 z_2 + 1}$$

or, somewhat simplifying the left-hand side of the last equation and transforming the right-hand side,

$$\arg\left(\frac{z_1 - z_2}{z_0 - z_2} : \frac{z_1 - z_3}{z_0 - z_3}\right) - \arg\left(\frac{\bar{z}_2 z_1 + 1}{\bar{z}_2 z_0 + 1} : \frac{\bar{z}_3 z_1 + 1}{\bar{z}_3 z_0 + 1}\right)$$
$$= \arg\left(\frac{z_0 - z_2}{z_1 - z_2} : \frac{z_0 - z_3}{z_1 - z_3}\right) + \arg\left(\frac{z_2 \bar{z}_1 + 1}{z_2 \bar{z}_0 + 1} : \frac{z_3 \bar{z}_1 + 1}{z_3 \bar{z}_0 + 1}\right)$$

But, because

$$\arg\left(\frac{z_1 - z_2}{z_0 - z_2} : \frac{z_1 - z_3}{z_0 - z_3}\right) = -\arg\left(\frac{z_0 - z_2}{z_1 - z_2} : \frac{z_0 - z_3}{z_1 - z_3}\right),$$

since

$$\frac{z_1 - z_2}{z_0 - z_2} : \frac{z_1 - z_3}{z_0 - z_3} = 1 : \left(\frac{z_0 - z_2}{z_1 - z_2} : \frac{z_0 - z_3}{z_1 - z_3}\right),$$

and

$$\arg\left(\frac{\bar{z}_2 z_1 + 1}{\bar{z}_2 z_0 + 1} : \frac{\bar{z}_3 z_1 + 1}{\bar{z}_3 z_0 + 1}\right) = -\arg\left(\frac{z_2 \bar{z}_1 + 1}{z_2 \bar{z}_0 + 1} : \frac{z_3 \bar{z}_1 + 1}{z_3 \bar{z}_0 + 1}\right),$$

since

$$\frac{\bar{z}_2 z_1 + 1}{\bar{z}_2 z_0 + 1} : \frac{\bar{z}_3 z_1 + 1}{\bar{z}_3 z_0 + 1} = \overline{\left(\frac{z_2 \bar{z}_1 + 1}{z_2 \bar{z}_0 + 1} : \frac{z_3 \bar{z}_1 + 1}{z_3 \bar{z}_0 + 1}\right)},$$

the condition, Equation 41, may be rewritten simply:

$$\arg\left(\frac{z_0 - z_2}{z_1 - z_2} : \frac{z_0 - z_3}{z_1 - z_3}\right) = 0 \tag{42}$$

The dual number $(z_0 - z_2)/(z_1 - z_2) : (z_0 - z_3)/(z_1 - z_3)$ is naturally called the **cross-ratio of the four lines** z_0, z_1, z_2, z_3; we denote it by the same symbol $W(z_0, z_1, z_2, z_3)$ as that used for the cross-ratio of four points (see p. 32). Thus, *the condition*

that four lines z_0, z_1, z_2, *and* z_3 *should touch one* (oriented) *circle* (circle of nonzero radius or circle of zero radius, a point) *is that the cross-ratio of these four lines should be real* (cf. p. 32).

The last condition may also be given as the familiar Equation 12

$$\frac{z_0 - z_2}{z_1 - z_2} : \frac{z_0 - z_3}{z_1 - z_3} = \frac{\bar{z}_0 - \bar{z}_2}{\bar{z}_1 - \bar{z}_2} : \frac{\bar{z}_0 - \bar{z}_3}{\bar{z}_1 - \bar{z}_3}$$

Hence it follows that the equation of an oriented circle (which in a special case may be a point) determined by three given points z_1, z_2, and z_3, has the form of Equation 13 (see Figure 34):[35]

$$\frac{z - z_2}{z_1 - z_2} : \frac{z - z_3}{z_1 - z_3} = \frac{\bar{z} - \bar{z}_2}{\bar{z}_1 - \bar{z}_2} : \frac{\bar{z} - \bar{z}_3}{\bar{z}_1 - \bar{z}_3}$$

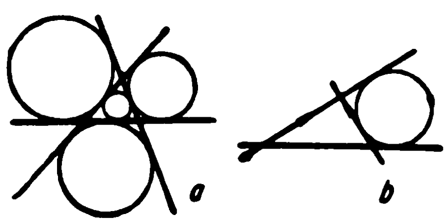


FIG. 34

Thus we see that the equation of *every* (oriented) *circle* (or point) *of the plane can be written in the form* (Equation 14):[36]

$$Az\bar{z} + Bz - \bar{B}\bar{z} + C = 0, \qquad A, C \text{ purely imaginary}$$

[35] We note that while, generally speaking, three given nonoriented lines are touched by four different nonoriented circles (Figure 34a), three oriented lines determine a *unique* oriented circle (Figure 34b). The latter statement only means that three given oriented lines cannot be touched by two different oriented circles; if two of the three lines are parallel, then there is *no* oriented circle which touches our three lines, and in this case the cross-ratio $W(z, z_1, z_2, z_3)$ cannot be real for any z.

[36] This result may be deduced from the fact that by the translation given by Equation 33 any circle S can be taken into the circle $\varepsilon r z\bar{z} - z + \bar{z} + \varepsilon r = 0$ with center at the origin O, r being the radius of the circle (cf. Equation 30). Hence it follows that the equation of S has the form

$$\varepsilon r \frac{p_1 z + q}{-\bar{q}z + \bar{p}_1} \cdot \frac{\bar{p}_1\bar{z} + \bar{q}}{-q\bar{z} + p_1} - \frac{p_1 z + q}{-\bar{q}z + \bar{p}_1} + \frac{\bar{p}_1\bar{z} + \bar{q}}{-q\bar{z} + p_1} + \varepsilon r = 0$$

It is not difficult to verify that, conversely, this *equation always represents a circle* (or point).[36a]

We already know that Equation 14 represents a *point* if

$$A + C = 0 \tag{43}$$

*§10. Applications and Examples

The theory developed above enables us to use dual numbers to prove numerous geometrical theorems relating to points, lines, and circles; in addition, the similarity between the results of Section 9 and those of Section 7 sometimes enables us to use *the same* calculation to prove two different propositions; it is sufficient to regard the numbers which occur in the reasoning as ordinary complex numbers in one case and dual numbers in the other case. We confine ourselves to a few examples, which illustrate this.

We begin with the following theorem. *Let four* (oriented) *circles S_1, S_2, S_3, and S_4 be given in a plane; let z_1 and w_1 be the common tangents of S_1 and S_2, z_2 and w_2 the common tangents of S_2 and S_3, z_3 and w_3 the common tangents of S_3 and S_4, and z_4 and w_4 the common tangents of S_4 and S_1. If the lines z_1, z_2, z_3, and z_4 touch one* (oriented) *circle Σ or pass through one point, then the lines w_1, w_2, w_3, and w_4 touch one* (oriented) *circle Σ' or pass through one point*; see Figure 35.

[It may be noted here that the "tangents" to an oriented circle S are naturally those lines which belong to S, regarded as a "locus of lines." It is obvious that, while two ordinary (non-oriented) circles have, generally speaking, four common tangents, two oriented circles S_1 and S_2 cannot have more than two common tangents. These are external common tangents, if S_1 and S_2

---

that is, the form of Equation 14, where $A = \varepsilon r p_1 \bar{p}_1 + p_1 q - \bar{p}_1 \bar{q} + \varepsilon r q \bar{q} = \varepsilon(r + t_1)$, $B = -p_1^2 - \bar{q}^2 = -(1 + \varepsilon t)$, and $C = \varepsilon r q \bar{q} - q p_1 + \bar{q} \bar{p}_1 + \varepsilon r p_1 \bar{p}_1 = \varepsilon(r - t_1)$, and where $-t$ and $-t_1$ are the coordinates of the center of S (see Footnote 30).

[36a] Among "circles" we should include pencils of parallel lines (the equation of any such pencil can be written in the form $Az\bar{z} + C = 0$, where A and C are purely imaginary).

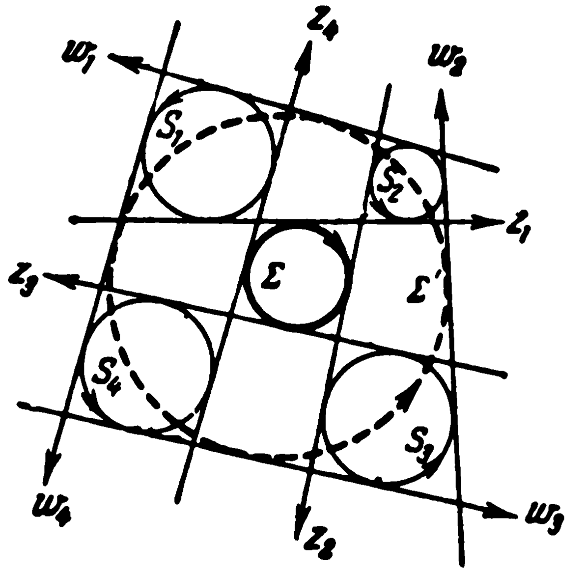


Fig. 35

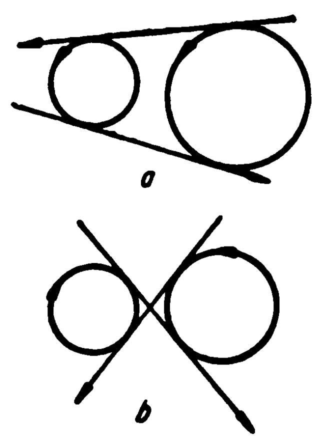


Fig. 36

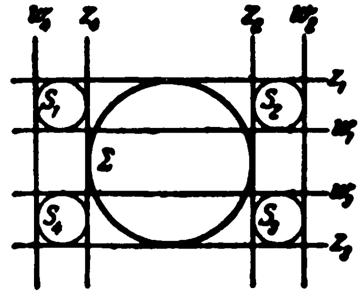


Fig. 37

have the same direction, and internal common tangents otherwise; see Figure 36. We note that the stated theorem becomes false if the circles S_1, S_2, S_3, S_4, and Σ are regarded as *nonoriented*; thus, in Figure 37 the lines w_1, w_2, w_3, and w_4 obviously do not touch one circle.]

From the conditions of the theorem it follows that the cross-ratios $W(z_1, z_2, w_1, w_2)$, $W(z_2, z_3, w_2, w_3)$, $W(z_3, z_4, w_3, w_4)$, and $W(z_4, z_1, w_4, w_1)$ are real, and so is the cross-ratio $W(z_1, z_3, z_2, z_4)$. It must be shown that the cross-ratio $W(w_1, w_3, w_2, w_4)$ also is real (see p. 94). The proof of this was given at the beginning of Section 8; the reasoning proves both the theorem given there and the present theorem.

We now turn to an arbitrary circle S (Equation 14):

$$Az\bar{z} + Bz - \bar{B}\bar{z} + C = 0, \qquad A,\ C \text{ purely imaginary}$$

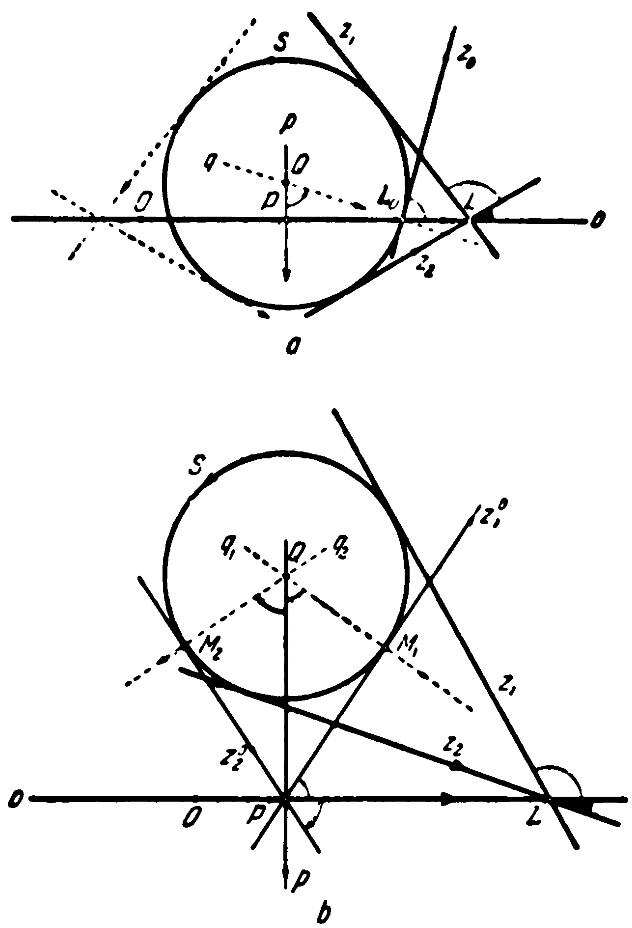


FIG. 38

We consider any point L of the axis o which is outside S; two tangents z_1 and z_2 to the circle S pass through this point; see Figure 38. Our problem consists of determining the value of the product (see Section 8, pp. 42–44):

$$\tan\frac{\angle\{o, z_1\}}{2}\cdot\tan\frac{\angle\{o, z_2\}}{2}$$

Since the lines z_1 and z_2 are tangents to the circle S, they satisfy Equation 14, so Equations 16 and 17 hold:

$$Az_1\bar{z}_1 + Bz_1 - \bar{B}\bar{z}_1 + C = 0$$

$$Az_1\bar{z}_2 + Bz_2 - \bar{B}\bar{z}_2 + C = 0$$

But from the definition of the argument of a dual number z corresponding to a definite line of the plane (see Figure 28) it follows that $\arg z_2 = \arg z_1$ and $\arg\bar{z}_2 = -\arg z_1$. Hence it follows that the product $z_1\bar{z}_2$ is a real number (Equations 18):

$$z_1\bar{z}_2 = k, \qquad \bar{z}_1 z_2 = \bar{k} = k$$

Multiplying Equation 16 by z_2 and Equation 17 by z_1 and using Equations 18, we obtain Equations 16a and 17a:

$$Akz_1 + Bz_1z_2 - \bar{B}k + Cz_2 = 0$$

$$Akz_2 + Bz_1z_2 - \bar{B}k + Cz_1 = 0$$

Subtracting the second of those equations from the first, we have

$$Ak(z_1 - z_2) - C(z_1 - z_2) = 0$$

whence it follows that, if $z_1 - z_2 \neq 0$ (that is, if $z_2 \neq z_1$), then

$$k = \frac{C}{A}$$

We now note that

$$z_1\bar{z}_2 = \tan\frac{\angle\{o, z_1\}}{2}(1+\varepsilon s)\cdot\tan\frac{\angle\{o, z_2\}}{2}(1-\varepsilon s)$$

$$= \tan\frac{\angle\{o, z_1\}}{2}\tan\frac{\angle\{o, z_2\}}{2}$$

Thus we have

$$\tan \frac{\angle\{o, z_1\}}{2} \tan \frac{\angle\{o, z_2\}}{2} = k = \frac{C}{A} \tag{44}$$

The quantity $C/A = k$ is called the **power of the circle** S (more precisely, *the power of the line* o *with respect to* S);[37] its geometrical meaning is given by Equation 44 (thus, *the product* $\tan \angle\{o, z_1\}/2 \cdot \tan \angle\{o, z_2\}/2$ *does not depend on the choice of the point* L *on the axis* o). We have obtained Equation 44 on the assumption that the tangents z_1 and z_2 of the circle S are distinct. But, if L_0 is a point of intersection of the axis o and the circle S, through which a unique tangent z_0 to the circle can be drawn (Figure 38a), and if $z_1 = z_2 = z_0$, then the product $\tan \angle\{o, z_1\}/2 \cdot \tan \angle\{o, z_2\}/2 = \tan^2 \angle\{o, z_0\}/2$ is also equal to C/A; it follows from this that the quantity $\tan^2 \angle\{o, z_0\}/2$ can be regarded as the limit of the expression $\tan \angle\{o, z_1\}/2 \cdot \tan \angle\{o, z_2\}/2$, where z_1 and z_2 are the tangents to the circle S, which pass through a variable point on the line o, which point approaches L_0. Thus, *the power of a line* o *with respect to a circle* S *which intersects it is equal to the square of the tangent of half the angle* $\angle\{o, S\}$ *which* o *makes with* S (since the angle $\angle\{o, S\}$ between an oriented line o and an oriented circle S which intersects it is understood to be the angle between o and the tangent to S at its point of intersection with o).

We now denote by P the foot of the perpendicular p, dropped from the center Q of the circle S to the line o, and put $\{Q, o\} = \{Q, P\} = d$ (these equations also define the oriented line p) and $\{Q, z_1\} = \{Q, z_2\} = r$; see Figure 38. In the case in which o intersects S we denote by q the line $[Q, L_0]$, whose orientation is determined by the condition $\{Q, L_0\} = r$; see Figure 38a. It is

[37] The quantity $k = C/A$ depends not only on the circle, Equation 14, but also on the choice of the coordinate system. It is not difficult to see, however, that k depends only on the position of the polar axis o and not on the choice of the pole O on this axis. This follows from the fact that, for any translation along o, given by $z' = pz$ and $z = p'z'$, where $|p| = 1$ and $|p'| = |1/p| = 1$ (see Equation 32), the circle, Equation 14, goes into the circle $Ap'\bar{p}'z'\bar{z}' + Bp'z' - \bar{B}\bar{p}'\bar{z}' + C = 0$, which has the same power $k = C/A$ (since $p'\bar{p}' = 1$).

not difficult to see that $\angle\{o, z_0\} = \angle\{p, q\}$, since p and q are perpendicular to o and z_0, respectively, and that $\cos \angle\{p, q\} = d/r$, whence it follows that

$$k = \tan^2 \frac{\angle\{o, z_0\}}{2} = \tan^2 \frac{\angle\{p, q\}}{2} = \frac{1 - \cos \angle\{p, q\}}{1 + \cos \angle\{p, q\}}$$

$$= \frac{1 - d/r}{1 + d/r} = \frac{r - d}{r + d}$$

On the other hand, if o does not intersect S (Figure 38b), then two tangents z_1^0 and z_2^0 to the circle S pass through P; the perpendiculars dropped from Q to z_1^0 and z_2^0 and oriented so that $\{Q, M_1\} = \{Q, M_2\} = r$ (where M_1 and M_2 are the points of contact of z_1 and z_2 with the circle S) are denoted by q_1 and q_2. In this case we have $\angle\{o, z_1^0\} = \angle\{p, q_1\}$ and $\angle\{o, z_2^0\} = \angle\{p, q_2\}$, since p, q_1, and q_2 are perpendicular to o, z_1^0, and z_2^0, respectively, and we have $\cos \angle\{p, q_1\} = \cos \angle\{p, q_2\} = r/d$. Since, moreover, it is obvious that $\angle\{o, z_2^0\} = -\angle\{o, z_1^0\}$ and that $\tan \angle\{o, z_2^0\}/2 = -\tan \angle\{o, z_1^0\}/2$, we have

$$k = \tan \frac{\angle\{o, z_1^0\}}{2} \cdot \tan \frac{\angle\{o, z_2^0\}}{2} = -\tan^2 \frac{\angle\{o, z_1^0\}}{2}$$

$$= -\tan^2 \frac{\angle\{p, q_1\}}{2} = -\frac{1 - \cos \angle\{p, q_1\}}{1 + \cos \angle\{p, q_1\}}$$

$$= -\frac{1 - r/d}{1 + r/d} = \frac{r - d}{r + d}$$

Thus, *in all cases,*

$$k = \frac{r - d}{r + d} \tag{45}$$

From Equation 45 it follows that the power of a line o with respect to the circle S is positive if o intersects S, is equal to zero if o touches S, becomes infinite if o **antitouches** S (that is,

if S touches the line o_1, which differs from o only in direction), and is negative if o does not intersect S (in the last case *the power of o with respect to S is equal to* $-\tan^2 \angle\{z_1^0, z_2^0\}/4$, where $\angle\{z_1^0, z_2^0\} = \varphi$ is the angle subtended by the circle S at the point P, the foot of the perpendicular dropped from Q to o).

The concept of the power of a line with respect to a circle enables us to give a geometrical meaning to the vanishing of the coefficients A or C in the equation of a circle S (Equation 14). If $C = 0$ and $A \neq 0$, the power of the line o with respect to S (Equation 45) is equal to zero, and so o touches S. If $A = 0$ and $C \neq 0$, the power of o with respect to S becomes infinite; that is, o antitouches S. Finally, the condition $A = C = 0$ means that o simultaneously touches and antitouches S, that is, that the line o and the line o_1, which differs from it in direction, are both tangents to S, but this is possible only if S is a point of the line o (cf. the condition, Equation 43, that Equation 14 should represent a point).

We now define the *power k of an* (arbitrary, oriented) *line w with respect to an* (oriented) *circle S* as the product

$$\tan \frac{\angle\{w, z_1\}}{2} \cdot \tan \frac{\angle\{w, z_2\}}{2}$$

where z_1 and z_2 are any two tangents to the circle S which meet in a point L of the line w; see Figure 39. From what was said

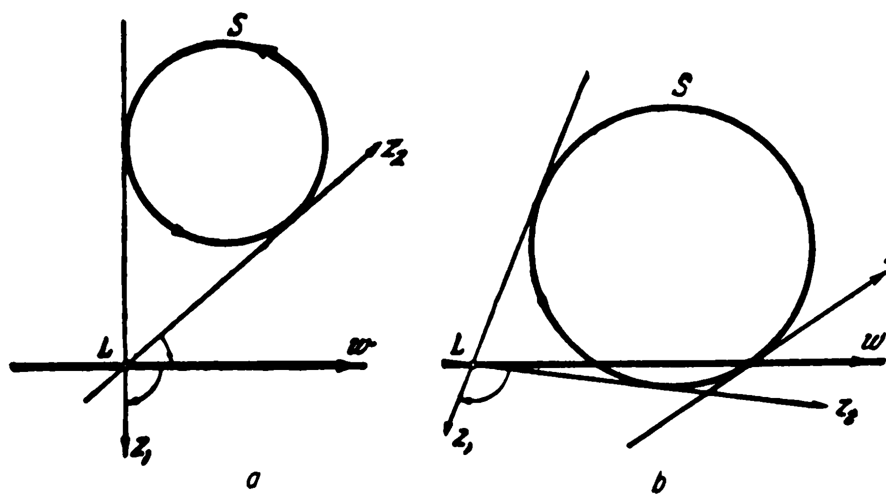


FIG. 39

above it follows that, if w intersects S, then the power of w with respect to S is equal to $\tan^2 \angle\{w, S\}/2$, where the angle $\angle\{w, S\}$ between the line w and the circle S is understood to be the angle $\angle\{w, z_0\}$ between w and the tangent z_0 to S at the point of intersection of w and S; if d is the (positive or negative) distance of the center of S from w, and r is the (positive or negative) radius of S, then the power of w with respect to S is equal to $(r - d)/(r + d)$. In particular, the power of w with respect to S is positive if w intersects S and negative if w does not intersect S; it is equal to zero if w touches S, and becomes infinite if w antitouches S.

Let us calculate the power of an arbitrary line w, where w is some dual number $w = \tan \varphi/2 \cdot (1 + \varepsilon s)$, with respect to the circle, Equation 14. We note that if a new system of dual line coordinates is introduced,[38]

$$Z = \frac{z - w}{\bar{w}z + 1}, \qquad z = \frac{Z + w}{-\bar{w}Z + 1} \tag{46}$$

then the role of the axis o of this system of coordinates is carried out by our line w (since, if $Z = 0$, it follows that $z = w$). The equation of the circle, Equation 14, in the new system of coordinates has the form

$$A\frac{Z + w}{-\bar{w}Z + 1} \cdot \frac{\bar{Z} + \bar{w}}{-w\bar{Z} + 1} + B\frac{Z + w}{-\bar{w}Z + 1} - \bar{B}\frac{\bar{Z} + \bar{w}}{-w\bar{Z} + 1} + C = 0$$

---

[38] The line z', which corresponds in the *new coordinate system* to the dual number z, is obtained from the line z, which corresponds to the same number in the *old coordinate system*, by the motion $z' = (z + w)/(-\bar{w}z + 1)$, which is represented by the translation $z_1 = z(1 - \varepsilon s)$ along o through a distance $-s$ together with the rotation $z_2 = (z_1 + \tan \phi/2)/(-\tan \phi/2 \cdot z_1 + 1)$ about O through an angle ϕ and a further translation $z' = z_2(1 + \varepsilon s)$ in the direction of o but through a distance $+s$. It is sometimes said that *the coordinate system Z is obtained from the coordinate system z by this motion.*

or

$$(A - Bw + \bar{B}\bar{w} + Cw\bar{w})Z\bar{Z} + [(A - C)\bar{w} + B + \bar{B}w\bar{w}]Z$$
$$- [(C - A)w + \bar{B} + Bw\bar{w}]\bar{Z}$$
$$+ (Aw\bar{w} + Bw - \bar{B}\bar{w} + C) = 0$$

Since the line w is the axis of the new system of coordinates, *the power of the line w with respect to the circle* (Equation 14) *is equal to*

$$k = \frac{Aw\bar{w} + Bw - \bar{B}\bar{w} + C}{A - Bw + \bar{B}\bar{w} + Cw\bar{w}} \tag{47}$$

because the power of the axis o with respect to the circle, Equation 14, is equal to C/A. Moreover, it follows from Equation 47 that

$$\frac{k-1}{k+1} = \frac{(A - C)w\bar{w} + 2Bw - 2\bar{B}\bar{w} + (C - A)}{(A + C)(w\bar{w} + 1)} \tag{47a}$$

From Equation 47 it follows immediately that, for example, *all lines w, which have a given power k with respect to a definite circle* (Equation 14) satisfy the equation:

$$\frac{Aw\bar{w} + Bw - \bar{B}\bar{w} + C}{A - Bw + \bar{B}\bar{w} + Cw\bar{w}} = k$$

which may be written

$$(A - kC)w\bar{w} + (k + 1)Bw - (k + 1)\bar{B}\bar{w} + (C - kA) = 0$$

That is, they are tangents to some circle (which, as is easily seen, is concentric with the original circle). Further, we consider two circles S_1 and S_2 with equations

$$A_1w\bar{w} + B_1w - \bar{B}_1\bar{w} + C_1 = 0$$

and

$$A_2w\bar{w} + B_2w - \bar{B}_2\bar{w} + C_2 = 0$$

where, to simplify the subsequent calculation, we shall assume that the sums of the first and last coefficients are equal:

$$A_1 + C_1 = A_2 + C_2$$

If this condition does not hold originally, we can always satisfy it by multiplying one of the equations by a suitably chosen real number. From Equation 47a it follows that *all lines w, whose powers with respect to S_1 and S_2 are equal,* satisfy the equation

$$\frac{(A_1 - C_1)w\bar{w} + 2B_1 w - 2\bar{B}_1\bar{w} + (C_1 - A_1)}{(A_1 + C_1)(w\bar{w} + 1)}$$
$$= \frac{(A_2 - C_2)w\bar{w} + 2B_2 w - 2\bar{B}_2\bar{w} + (C_2 - A_2)}{(A_2 + C_2)(w\bar{w} + 1)},$$

that is, they satisfy the following equation, since, by virtue of our condition, $A_1 - A_2 = -(C_1 - C_2)$ or $(A_1 - A_2) - (C_1 - C_2) = 2(A_1 - A_2)$:

$$(A_1 - A_2)w\bar{w} + (B_1 - B_2)w - (\bar{B}_1 - \bar{B}_2)\bar{w} - (A_1 - A_2) = 0$$

That is, all these lines *pass through one point Q* (cf. Equation 40, the equation of a point). If the circles S_1 and S_2 have two common tangents, then Q coincides with their point of intersection (since both tangents have zero power with respect to S_1 and with respect to S_2); see Figure 40a. In the general case the point Q is characterized by the fact that all lines which pass through it intersect S_1 and S_2 at the same angle (the square of the tangent of half of each is equal to the power of the line with respect to S_1 and S_2). The point Q is called the **center of similitude** of S_1 and S_2; see Figure 40).[39]

Finally, we consider three circles S_1, S_2, and S_3 with equations

$$A_1 z\bar{z} + B_1 z - \bar{B}_1\bar{z} + C_1 = 0$$
$$A_2 z\bar{z} + B_2 z - \bar{B}_2\bar{z} + C_2 = 0$$
$$A_3 z\bar{z} + B_3 z - \bar{B}_3\bar{z} + C_3 = 0$$

where, as above, we assume that $A_1 + C_1 = A_2 + C_2 = A_3 + C_3$.

---

[39] It is not difficult to verify that this point is the center of the dilatation (homothety) which takes S_1 into S_2; the ratio of this transformation is equal to the ratio r_2/r_1 of the radii of the circles. See H. S. M. Coxeter, *Introduction to Geometry*, §5.1 (John Wiley, New York), 1961 and I. M. Yaglom, *Geometric Transformations*, Part II, Chap. 1, Section 1 (Random House, New York).

The centers of similitude of the circles S_1, S_2, and S_3 in pairs are characterized by the equations

$$(A_1 - A_2)w\bar{w} + (B_1 - B_2)w - (\bar{B}_1 - \bar{B}_2)\bar{w} - (A_1 - A_2) = 0$$
$$(A_1 - A_3)w\bar{w} + (B_1 - B_3)w - (\bar{B}_1 - \bar{B}_3)\bar{w} - (A_1 - A_3) = 0$$
$$(A_2 - A_3)w\bar{w} + (B_2 - B_3)w - (\bar{B}_2 - \bar{B}_3)\bar{w} - (A_2 - A_3) = 0$$

From these equations it follows that the line which passes through the first two centers of similitude also passes through the third

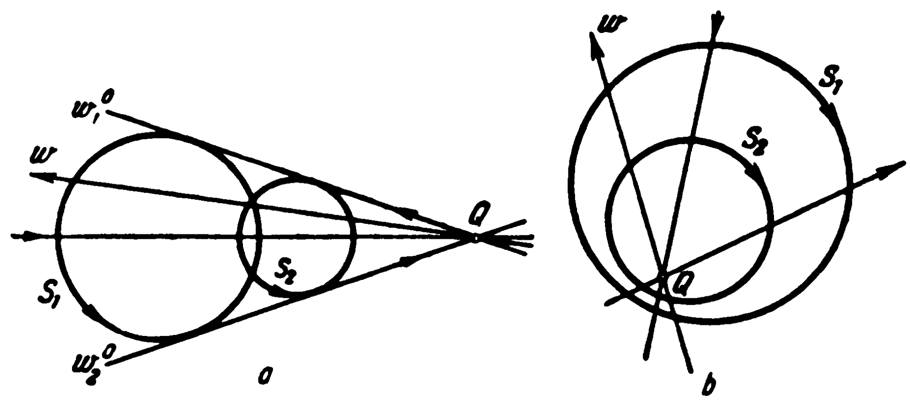


FIG. 40

(since the last of the three equations is the difference of the first two, and so any line w which satisfies the first two equations necessarily satisfies the third). Thus we see that *the centers of simulitude of three circles S_1, S_2, and S_3 in pairs lie on one line q.* This line is called the **axis of similitude** of the three circles S_1, S_2, and S_3; see Figure 41.

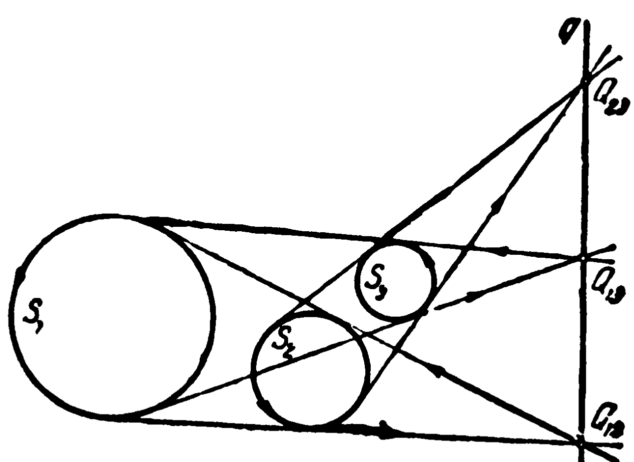


FIG. 41

We turn now to the idea of the *cross-ratio* of four lines z_1, z_2, z_3, and z_4. It is clear that a permutation of these lines may alter the cross-ratio W; in fact, it follows from the definition of cross-ratio that *W changes to its reciprocal if the first two lines or the last two lines are interchanged, does not change if the first pair of lines is interchanged with the second pair*, and *changes to its complement with respect to unity if the second and third lines are interchanged* (cf. p. 72). Hence, exactly as before, we conclude that *for all possible permutations of the four lines we obtain six different values of the cross-ratio* (see Equation 26):

$$\lambda, \qquad \lambda_1 = \frac{1}{\lambda}, \qquad \lambda_2 = 1 - \lambda,$$
$$\lambda_3 = \frac{1}{1-\lambda}, \qquad \lambda_4 = \frac{\lambda - 1}{\lambda}, \qquad \lambda_5 = \frac{\lambda}{\lambda - 1}$$

Thus, to each quadrilateral $\overline{z_1 z_2 z_3 z_4}$ (that is, a set of four lines z_1, z_2, z_3, and z_4, the sides of the quadrilateral) there corresponds a definite hexagon,[40] $\overline{\lambda\lambda_1\lambda_2\lambda_3\lambda_4\lambda_5}$ given by the six lines (six dual numbers) λ, λ_1, λ_2, λ_3, λ_4, and λ_5.[41]

Let us determine when the hexagon $\overline{\lambda\lambda_1\lambda_2\lambda_3\lambda_4\lambda_5}$ will be degenerate, that is, will have less than six sides. To do this, we must find under what circumstances the (dual) number λ will be equal to one of the following five numbers: $\lambda_1 = 1/\lambda$, $\lambda_2 = 1 - \lambda$, $\lambda_3 = 1/(1 - \lambda)$, $\lambda_4 = (\lambda - 1)/\lambda$, and $\lambda_5 = \lambda/(\lambda - 1)$. But we know already that the equation $\lambda = \lambda_1$ leads to the value $\lambda = -1$ (we note that $W(z_1, z_2, z_3, z_4) = (z_1 - z_3)/(z_2 - z_3) : (z_1 - z_4)/(z_2 - z_4) \neq 1$, if the lines z_1, z_2, z_3 and z_4 are distinct); the equation $\lambda = \lambda_2$ leads to the values $\lambda = \frac{1}{2}$ and $\lambda_4 = -1$; from the equation $\lambda = \lambda_5$ it follows that $\lambda = 2$ and $\lambda_2 = -1$ (we recall that the equation $\lambda = (z_1 - z_3)/(z_2 - z_3) : (z_1 - z_4)/(z_2 - z_4) = 0$ is impossible, if the lines z_1, z_2, z_3 and z_4 are distinct); finally, from the equations $\lambda = \lambda_3$ and $\lambda = \lambda_4$ it

[40] Strictly speaking, we should use two different words for the figure formed by six points and that formed by six lines (just as we have "quadrangle" and "quadrilateral"); however, it is customary to use the word "hexagon" for both.—TRANSL.

[41] This hexagon is completely determined by giving one of its sides λ.

follows that $\lambda^2 - \lambda + 1 = 0$ (see pp. 70–71). But the equation $\lambda^2 - \lambda + 1 = 0$ is insoluble in the domain of dual numbers (cf. footnote 6); in fact, if $\lambda = a + b\varepsilon$, where a and b are real numbers, then

$$\begin{aligned}\lambda^2 - \lambda + 1 &= (a^2 + 2ab\varepsilon) - (a + b\varepsilon) + 1\\ &= (a^2 - a + 1) + (2ab - b)\varepsilon\end{aligned}$$

and the "real part" $a^2 - a + 1$ of the last number cannot equal zero for any (real!) a.

Thus, the only case in which the cross-ratio W of four (distinct) lines takes less than six values is one in which the lines, denoted by z_1, z_2, z_3, and z_4, satisfy the condition given by Equation 27:

$$W(z_1, z_2, z_3, z_4) = -1$$

Four such lines are called a **harmonic tetrad**, and the quadrilateral $\overline{z_1z_2z_3z_4}$ formed by these lines is called a **harmonic quadrilateral**; see Figure 42. Since the cross-ratio of the four

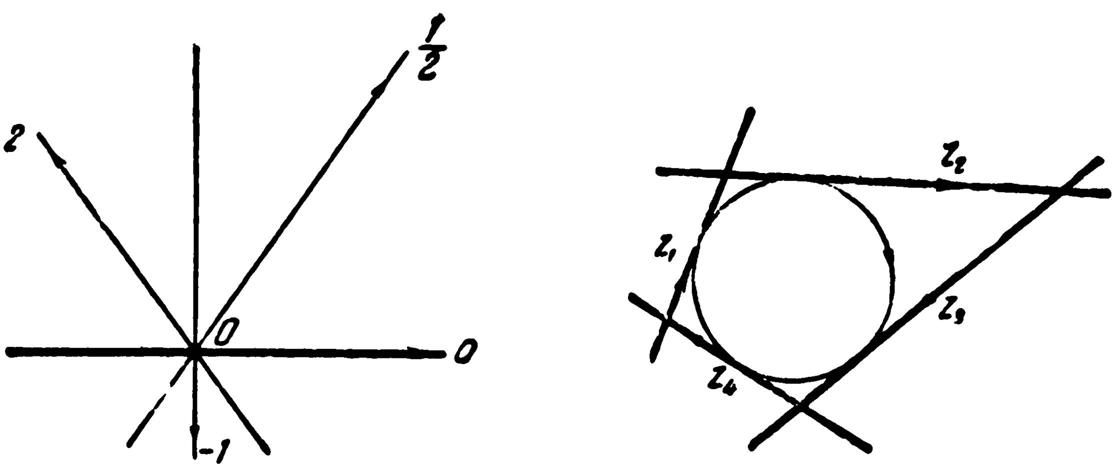


FIG. 42

sides of a harmonic quadrilateral is real, *a harmonic quadrilateral can be circumscribed about a circle.* On the other hand,

$$\begin{aligned}|W(z_1, z_2, z_3, z_4)| &= \frac{|z_1 - z_3|}{|z_2 - z_3|} : \frac{|z_1 - z_4|}{|z_2 - z_4|}\\ &= \frac{\tan(\varphi_1/2) - \tan(\varphi_3/2)}{\tan(\varphi_2/2) - \tan(\varphi_3/2)} : \frac{\tan(\varphi_1/2) - \tan(\varphi_4/2)}{\tan(\varphi_2/2) - \tan(\varphi_4/2)}\\ &= \frac{\sin(\varphi_1 - \varphi_3)/2}{\sin(\varphi_2 - \varphi_3)/2} : \frac{\sin(\varphi_1 - \varphi_4)/2}{\sin(\varphi_2 - \varphi_4)/2}\end{aligned}$$

where $\varphi_1 = \angle\{z_1, o\}$, $\varphi_2 = \angle\{z_2, o\}$, and so on (here we make use of the formula $\tan\alpha - \tan\beta = \sin(\alpha - \beta)/(\cos\alpha\cos\beta)$. And, since $\varphi_1 - \varphi_3 = \angle\{z_1, o\} - \angle\{z_3, o\} = \angle\{z_1, z_3\}$, $\varphi_1 - \varphi_4 = \angle\{z_1, o\} - \angle\{z_4, o\} = \angle\{z_1, z_4\}$ and so on, the equation $|W(z_1, z_2, z_3, z_4)| = 1$ can be rewritten as

$$\frac{\sin\dfrac{\angle\{z_1, z_3\}}{2}}{\sin\dfrac{\angle\{z_2, z_3\}}{2}} : \frac{\sin\dfrac{\angle\{z_1, z_4\}}{2}}{\sin\dfrac{\angle\{z_2, z_4\}}{2}} = 1$$

or

$$\sin\frac{\angle\{z_1, z_3\}}{2}\cdot\sin\frac{\angle\{z_2, z_4\}}{2} = \sin\frac{\angle\{z_2, z_3\}}{2}\cdot\sin\frac{\angle\{z_1, z_4\}}{2}$$

Thus, *the products of the sines of half the opposite angles of a harmonic quadrilateral* $\overline{z_1z_2z_3z_4}$, where we understand the angles to be the directed angles $\angle\{z_1, z_3\}$ and $\angle\{z_1, z_4\}$ and so on, *are equal.* It is easy to see that the last two conditions completely characterize harmonic quadrilaterals; therefore, one example of a harmonic quadrilateral is a square (whose sides are oriented so that they all touch one circle with its center at the center of the square).

The role of the hexagon $\overline{\lambda\lambda_1\lambda_2\lambda_3\lambda_4\lambda_5}$ corresponding to a harmonic quadrilateral $\overline{z_1z_2z_3z_4}$ is played by the three lines -1, 2, and $\frac{1}{2}$, which meet at the pole O of the coordinate system (Figure 42).

**§11. Interpretation of Ordinary Complex Numbers in the Lobachevskii Plane

It is well known that the points of the Lobachevskii plane can be represented as points in the interior of the unit circle Σ; the lines of the Lobachevskii plane are then represented by the diameters of the circle Σ and arcs of circles which are orthogonal to the circle Σ.[42] This representation was first considered by the

[42] See, for example, H. S. M. Coxeter, *Introduction to Geometry*, §16.2 (John Wiley, New York) 1961 and I. M. Yaglom, *Geometric Transformations*, Part III (to be published by Random House, New York).

outstanding French mathematician and physicist H. Poincaré (1854–1912) and is therefore called the "Poincaré model" of the Lobachevskii plane. The Poincaré model may also be considered as a representation of the Lobachevskii plane on the plane of a complex variable;[43] it enables us to set up a correspondence between (ordinary) complex numbers and points of the Lobachevskii plane. This correspondence is set up in the following way. To a **point of the Lobachevskii plane** with polar coordinates (r, φ) corresponds the complex number

$$z = \tanh \frac{r}{2} (\cos \varphi + i \sin \varphi) \tag{48}$$

or, in other words, the complex number $z = \rho(\cos \varphi + i \sin \varphi)$ is represented by the point of the Lobachevskii plane with polar coordinates (r, φ), where $r = 2 \tanh^{-1} \rho$, i.e. $\tanh r/2 = \rho$. The whole Lobachevskii plane is represented by the set of numbers z such that $|z|^2 = z\bar{z} < 1$, that is, by the set of points inside the unit circle (see Section 7).

In order to extend the correspondence between ordinary complex numbers and points of the Lobachevskii plane to *all* complex numbers we can proceed as we did in Section 9, constructing the representation of Euclidean lines by dual numbers. Namely, we agree to regard all points of the Lobachevskii plane as *oriented* (that is, provided with an indication of a definite direction of rotation about the point, taken as positive); on a diagram an oriented point (one with a prescribed direction of rotation) will be denoted by a short curved arrow (Figure 43). We shall regard the distance $d = (A, B)$ between two oriented points A and B of the Lobachevskii plane as equal

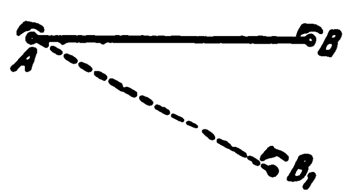


FIG. 43

[43] See, for example, H. Schwerdtfeger, *Geometry of Complex Numbers*, §14 (Toronto University Press, Oliver and Boyd), 1962.

to the (non-Euclidean) length r of the segment AB only if the positive directions of rotation about A and B coincide; if the orientations of the points A and B are different (Figure 43), we shall assume that the distance between these points is complex and equal to $r + i\pi$. If this is so, then according to Equation 48 two oriented points A and A_1 of the Lobachevskii plane with polar coordinates (r, φ), differing only in direction, correspond to complex numbers

$$z = \tanh\frac{r}{2}(\cos\varphi + i\sin\varphi)$$

$$\begin{aligned} z_1 &= \tanh\left(\frac{r}{2} + i\frac{\pi}{2}\right)(\cos\varphi + i\sin\varphi) \\ &= \coth\frac{r}{2}(\cos\varphi + i\sin\varphi) = \frac{1}{\bar{z}} \end{aligned}$$

The corresponding points of the complex plane are inverse with respect to the circle $z\bar{z} = 1$; these points lie on one ray with origin O, and $(O, z_1) = 1/(O, z)$, where (O, z) and (O, z_1) are the *Euclidean* distances from the point O to the points z and z_1 (later we shall have occasion to dwell at greater length on the idea of inversion with respect to a circle). In Figure 44 the complex

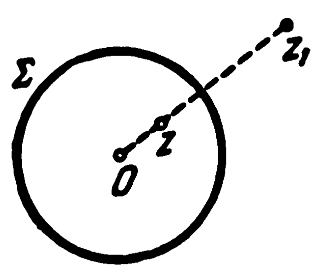


FIG. 44

number z corresponds to the point A, which is oriented in the same way as the pole O of the polar coordinate system, and the number z_1 corresponds to the point A_1, which differs from A only in orientation (the orientation of A_1 is opposite to the orientation of O). Clearly, the pole O and the **antipole** O_1 (the point differing from O only in orientation) correspond to the numbers 0 and ∞. If we agree to call the points of the circle $z\bar{z} = 1$ (the **absolute** of the Poincaré model) **points at infinity**

of the Lobachevskii plane[44] and assume that for these points the radius vector $r = \infty$, then we obtain a one-to-one correspondence between all the (oriented and infinite) points of the Lobachevskii plane and all the ordinary complex numbers (among which we include the number ∞).

Thus, ordinary complex numbers can be interpreted geometrically, not only as points of an ordinary (Euclidean) plane, but also as **oriented points of a Lobachevskii plane**. Here, as before, to real numbers correspond points of the polar axis o, to opposite numbers z and $-z$ correspond points symmetrical about the pole O, and to conjugate numbers z and $\bar{z}$ correspond points symmetrical about the polar axis o (see Figure 1); to points differing only in orientation correspond complex numbers z and z_1 such that $z_1 = 1/\bar{z}$. Thus, the equations

$$z' = -z, \qquad z' = \bar{z}, \qquad z' = \frac{1}{\bar{z}} \tag{49}$$

define in the Lobachevskii plane the **symmetry about the point** O, the **symmetry about the line** o, and the **reorientation** (change of orientation of each point to its opposite).

An arbitrary **motion** of the Lobachevskii plane is described by one of the formulae[45]

$$z' = \frac{pz + q}{\bar{q}z + \bar{p}} \quad \text{or} \quad z' = \frac{p\bar{z} + q}{\bar{q}\bar{z} + \bar{p}}, \qquad \Delta = \begin{vmatrix} p & q \\ \bar{q} & \bar{p} \end{vmatrix} \neq 0 \tag{50}$$

The **distance** d_0 from the pole O of the coordinate system to the point z is determined, by virtue of Equation 48, by the formula

$$\tanh \frac{d_0}{2} = |\, z \,|, \qquad \tanh^2 \frac{d_0}{2} = z\bar{z}$$

[44] We note that the points at infinity of the Lobachevskii plane *have no orientation;* this can be explained graphically by the fact that around such a point it is impossible to draw a circle, whose direction of rotation gives the orientation of the point.

[45] See L. Bieberbach, *Conformal Mapping*, pp. 36–38 (Chelsea, New York), 1964. Bieberbach is interested only in the first type of transformation and only in those which map the interior of the unit circle onto itself, so he has $\Delta > 0$. [A motion which maps the interior of the unit circle onto its exterior (a motion which changes the orientation of points) can be obtained as the product of a motion which maps the interior of the unit circle onto itself and the reorientation (Equation 49c).]

whence, using Equation 50, we easily find an expression for the distance $d(z_1, z_2)$ between any two points z_1 and z_2:

$$\tanh^2 \frac{d}{2} = \frac{(z_2 - z_1)(\bar{z}_2 - \bar{z}_1)}{(1 - z_1\bar{z}_2)(1 - \bar{z}_1 z_2)} \tag{51}$$

The **angle** δ between two lines which meet at O and pass through the points z_1^0 and z_2^0 is expressed, by virtue of Equation 48, by the same relation, Equation 7, as in the case of a Euclidean plane. Hence, using Equation 50, we may obtain the following expression for the (oriented) **angle** $\delta = \angle\{[z_0 z_1], [z_0 z_2]\}$ between the (oriented) lines $[z_0 z_1]$ and $[z_0 z_2]$:[46]

$$\delta = \arg\left(\frac{z_2 - z_0}{z_1 - z_0} : \frac{1 - \bar{z}_0 z_2}{1 - \bar{z}_0 z_1}\right) \tag{52}$$

By virtue of Equation 52 *the condition that the three points* z_0, z_1, *and* z_2 *should lie on one line is that the ratio* $(z_0 - z_2)/(z_1 - z_2) : (1 - \bar{z}_2 z_0)/(1 - \bar{z}_2 z_1)$ *should be real.* Hence it is not difficult to derive the *equation of a line* in the non-Euclidean geometry of Lobachevskii (see pp. 32–33 and 90):

$$Az\bar{z} + Bz - \bar{B}\bar{z} + A = 0, \qquad A \text{ purely imaginary} \tag{53}$$

By a line we understand the set of all *oriented* points lying on a given line.[47]

The **cycles** of Lobachevskii geometry comprise **circles**, **horocycles**, and **equidistant curves** (hypercycles); equidistant curves include lines, regarded as limiting cases of equidistant curves. It is well known that these cycles are represented by

[46] This formula clearly does not coincide with Equation 8, in spite of the fact that the "non-Euclidean angle" between two curves on the Poincaré model is represented by their ordinary (Euclidean) angle: this is because a "non-Euclidean line" $[z_0 z_1]$, which is, from the point of view of Euclidean geometry, a circle orthogonal to Σ, is different from an ordinary (Euclidean) line $[z_0 z_1]$.

[47] Equation 53 can also be deduced from the fact that by the motion given by Equation 50 any line of the Lobachevskii plane can be taken into the polar axis o, whose equation has the form $z - \bar{z} = 0$.

circles and lines of the plane of the complex variable.[48] This statement may be made somewhat more precise by explaining how the word "cycle" must be understood when the points of the Lobachevskii plane are regarded as oriented. It is natural to regard the cycles also as *oriented*, where an oriented point A is to be regarded as lying on an oriented cycle S, if the direction of rotation about A coincides with the direction of rotation for motion on the cycle; see the schematic Figure 45a, where the

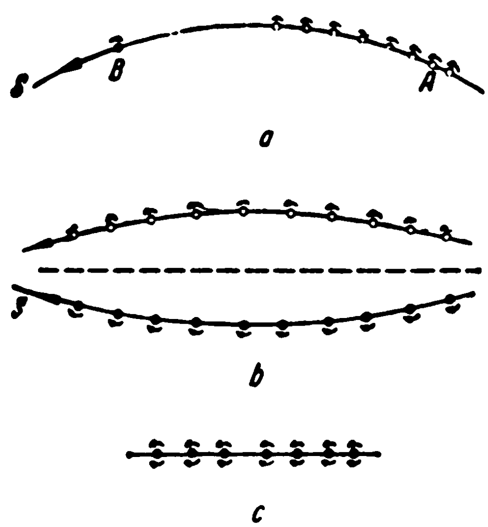


FIG. 45

point A lies on the cycle S, but the point B is not regarded as lying on it. Further, by an (oriented) equidistant curve with base PQ we understand the locus of points, whose distance h from the line PQ is constant and which lie on both sides of this line; the points on the upper and lower branches of the equidistant curve must be oriented differently (Figure 45b).[49] The lines of the Lobachevskii plane will be regarded as nonoriented (just as,

[48] See, for example, H. S. M. Coxeter, *Non-Euclidean Geometry,* §14.7 (Toronto University Press), 1957, and I. M. Yaglom, *Geometric Transformations*, Part III (to be published by Random House, New York).

[49] Since we are regarding a line as a limiting case of an equidistant curve, as h tends to zero, the points of the line should be regarded as *double*, and both possible orientations should be assigned to each of them (cf. Figure 45b, c); this means that in the plane of the complex variable a line represents a complete circle orthogonal to the absolute Σ (the circle $z\bar{z} = 1$).

in Section 9, points were regarded as nonoriented).[50] Finally, among the cycles we shall include the **circle at infinity** (the absolute) Σ, also nonoriented. Thus, the set of all (oriented) cycles of the Lobachevskii plane coincides exactly with the set of all circles and lines of the plane of the complex variable; see Figure 46, which shows a line of the Lobachevskii plane and an equidistant curve S, for which this line is the base.

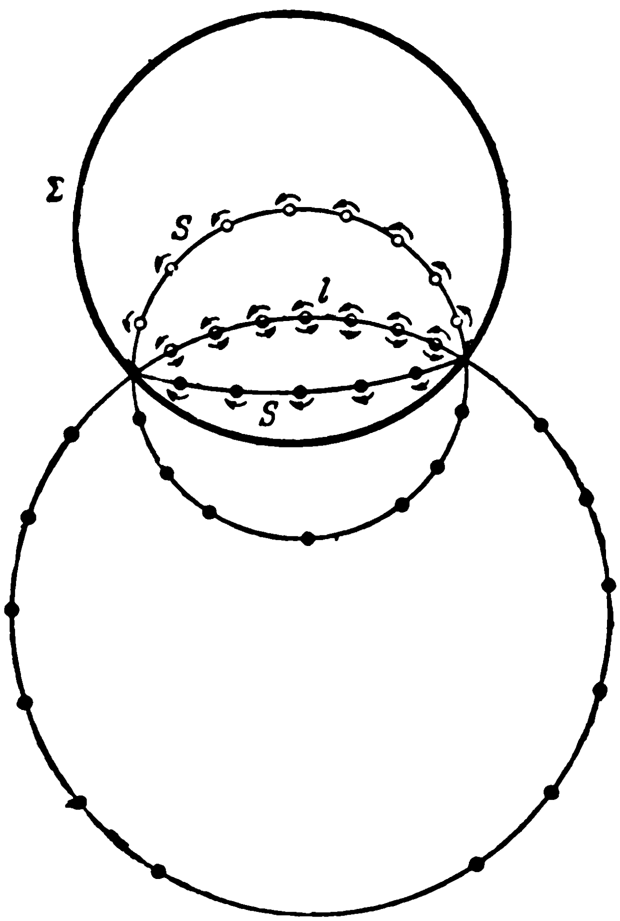


FIG. 46

[50] The idea of an oriented cycle was introduced in order to establish the orientation of points lying on the cycle: the direction of the arrow on the arc around the point, drawn *on the convex side* of the cycle, should coincide with the direction of the cycle (Figure 45a). However, since a line has no convexity, we cannot select a direction for the points lying on it (see Figure 45a); therefore, we must regard lines as undirected and all the points lying on them as double.

The last property enables us to use here the results of Section 7. Remembering, in particular, the condition that four points of the plane of the complex variable should lie on one circle (pp. 32–33), we conclude that *the condition that four given* (oriented) *points* z_0, z_1, z_2, *and* z_3 *of the Lobachevskii plane should lie on one* (oriented) *cycle is that the cross-ratio* $W(z_0, z_1, z_2, z_3)$ *of these points should be real.* Hence it follows that the *equation of every cycle of the Lobachevskii plane* can be written in the form (Equation 14):

$$Az\bar{z} + Bz - \bar{B}\bar{z} + C = 0, \qquad A,\ C \text{ purely imaginary}$$

Equation 14 will represent a circle, horocycle, or equidistant curve, depending on whether the circle (Equation 14) in the complex plane has 0, 1, or 2 points of intersection with the circle $z\bar{z} = 1$ (the absolute), that is, depending on the number of solutions of the system of equations

$$z\bar{z} = 1, \qquad Bz - \bar{B}\bar{z} = -A - C$$

Hence, with no difficulty, the following results are obtained.

The cycle, Equation 14, *is a circle* (among circles we include the circle at infinity Σ), *if*

$$AC + B\bar{B} > 0, \qquad (A + C)^2 + 4B\bar{B} < 0 \tag{54a}$$

It is a horocycle, if

$$AC + B\bar{B} > 0, \qquad (A + C)^2 + 4B\bar{B} = 0 \tag{54b}$$

It is an equidistant curve (among equidistant curves we include lines),[51] *if*

$$AC + B\bar{B} > 0, \qquad (A + C)^2 + 4B\bar{B} > 0 \tag{54c}$$

We have already seen that Equation 14 *represents a line, if*

$$A - C = 0 \tag{55}$$

[51] If $AC + B\bar{B} = 0$, then Equation 14 gives a unique point; if $AC + B\bar{B} < 0$, then Equation 14 is not satisfied by any point of the Lobachevskii plane.

From what has been said we can derive the proofs of many theorems of the non-Euclidean geometry of Lobachevskii. Thus, for example (see pp. 35–36), it can be shown that, *if* S_1, S_2, S_3, and S_4 *are four* (oriented) *cycles of the Lobachevskii plane, where the cycles* S_1 *and* S_2 *intersect at the* (oriented) *points* z_1 *and* w_1, *the cycles* S_2 *and* S_3 *intersect at the points* z_2 *and* w_2, *the cycles* S_3 *and* S_4 *intersect at the points* z_3 *and* w_3, *and the cycles* S_4 *and* S_1 *intersect at the points* z_4 *and* w_4, *and if the* (oriented) *points* z_1, z_2, z_3 *and* z_4 *lie on one* (oriented) *cycle, then the points* w_1, w_2, w_3, *and* w_4 *lie on one cycle.*[52]

The reader may discover other examples of this kind for himself.

The Poincaré model of Lobachevskii plane geometry is often understood to be a representation of the Lobachevskii plane somewhat different from that used above. In it the points of the Lobachevskii plane are represented by all the points of some half-plane, excluding the points of the line o which bounds this half-plane, and the lines are rays and semicircles orthogonal to o (in other words, rays perpendicular to o and semicircles with centers on o; see Figure 47).[53] The non-Euclidean

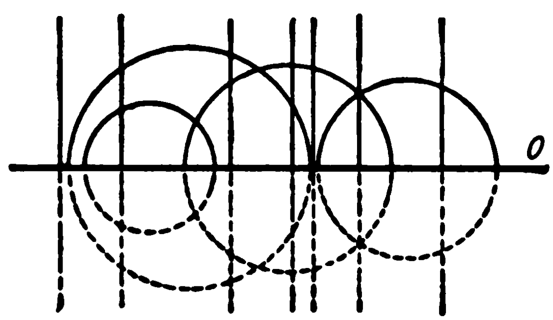


FIG. 47

distance between two points with complex coordinates $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$ (where $y_1 > 0, y_2 > 0$, since the role of points of the Lobachevskii plane is played here by points of the upper half-plane $y > 0$) is defined by the formula

$$\tanh^2 \frac{d}{2} = \frac{(z_2 - z_1)(\bar{z}_2 - \bar{z}_1)}{(\bar{z}_2 - z_1)(z_2 - \bar{z}_1)} \tag{56}$$

The non-Euclidean angle between two lines is measured by the

[52] We note, however, that this theorem follows automatically from the theorem proved on pp. 34–35 and from the fact that cycles of the non-Euclidean geometry of Lobachevskii are represented by circles of the Euclidean plane.

[53] See, for example, H. S. M. Coxeter, *Introduction to Geometry*, §16.7 (John Wiley, New York), 1961.

Euclidean angle between the circles (or line and circle) which represent these lines.

Let us now agree, as before, to regard the points of the Lobachevskii plane as *oriented*; with two points which differ only in direction we associate two points z and $\bar{z}$ of the complex plane, symmetrical about the axis o. If, in addition, the points of the line o (the *absolute* of this "Poincaré half-plane model") are called *points at infinity* of the Lobachevskii plane (these points have no orientation), then we again obtain a *representation of the whole plane of the complex variable* (the whole set of complex numbers) *on the set of* (oriented and infinite) *points of the Lobachevskii plane*. Assuming, as above, the condition about the orientation of cycles and about (oriented) points lying on (oriented) cycles, we also obtain the result that *the set of all cycles of the Lobachevskii plane is represented by the set of lines and circles of the plane of the complex variable*. Hence it follows that, as before, *the condition that four* (oriented) *points* z_1, z_2, z_3, *and* z_4 *should lie on one cycle is that the cross-ratio* $W(z_1, z_2, z_3, z_4)$ *of these four points should be real*, and that *the equation of an* (oriented) *cycle has the familiar form* (Equation 14):

$$Az\bar{z} + Bz - \bar{B}\bar{z} + C = 0, \qquad A, C \text{ purely imaginary}$$

Further, since the cycle, Equation 14, is an equidistant curve, a horocycle, or a circle, depending on whether it has two, one, or no common points with the absolute $z - \bar{z} = 0$ (the real axis o), we obtain with no difficulty the result that it *is a circle if* $(B - \bar{B})^2 - 4AC < 0$, *a horocycle if* $(B - \bar{B})^2 - 4AC = 0$, *and an equidistant curve if* $(B - \bar{B})^2 - 4AC > 0$.

It is not difficult to verify also that *the cycle*, Equation 14, *is a line if and only if*

$$B + \bar{B} = 0 \tag{57}$$

Finally we note that, in terms of the complex point coordinates considered here, *the motions of the Lobachevskii plane can be written as*

$$z' = \frac{az + b}{cz + d}, \qquad ad - bc > 0$$

or

$$z = \frac{a\bar{z} + b}{c\bar{z} + d}, \qquad ad - bc < 0 \tag{58}$$

where a, b, c, d are *real* numbers.

We shall return later (Section 17) to the question of the connection between these two different Poincaré models of the Lobachevskii plane (two representations of the Lobachevskii plane on the plane of the complex variable).

**§12. Double Numbers as Oriented Lines of the Lobachevskii Plane

In complete analogy with the results of Section 9, oriented lines of the Lobachevskii plane can be associated with double numbers. We introduce, as in Section 9, a system of polar coordinates for lines, and we associate with each line l which *intersects* the polar axis o and has polar coordinates (θ, s) a double number

$$z = \tan\frac{\theta}{2}(\cosh s + e \sinh s) \tag{59}$$

and with a line m, **ultraparallel** to (divergent from) o and directed in the same sense as o from their common perpendicular, the number

$$z = \tanh\frac{d}{2}(\sinh s' + e \cosh s') \tag{59a}$$

Here $d = \{m, o\} = \{P, Q\}$ is the shortest (oriented) distance between the lines m and o, that is, the oriented distance from o of the projection P on the line m of the common perpendicular of the lines m and o (see p. 80), and $s' = \{O, Q\}$ is the (oriented) distance from the pole O of the coordinate system to the projection Q of the common perpendicular on o; see Figure 48.[54] Further,

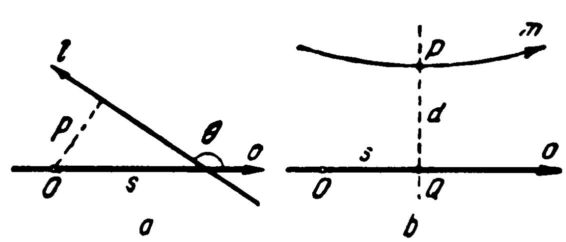


FIG. 48

[54] Equation 59a means that, in accordance with the general formulae of the non-Euclidean geometry of Lobachevskii, the polar coordinates of the line m are regarded as equal to $\theta = id$ and $s = s' - i(\pi/2)$, since $\tan(id/2) \cdot \{\cosh[s' - i(\pi/2)] + e \sinh[s' - i(\pi/2)]\} = \tanh(d/2) \cdot (\sinh s' + e \cosh s')$.

since it follows from Equation 59 that two lines l and l_1 which intersect o and differ only in direction correspond to the double numbers

$$z = \tan\frac{\theta}{2}(\cosh s + e\sinh s)$$

$$z_1 = \tan\frac{\theta+\pi}{2}(\cosh s + e\sinh s)$$

$$= -\cot\frac{\theta}{2}(\cosh s + e\sinh s) = -\frac{1}{\bar{z}},$$

then the line m_1, which differs only in direction from the line m, ultraparallel to o and corresponding to the number given by Equation 59a, is associated with the number

$$z = -\frac{1}{\tanh\frac{d}{2}(\sinh s' + e\cosh s')} = \coth\frac{d}{2}(\sinh s' + e\cosh s') \tag{59b}$$

Lines *parallel* to the axis o can be regarded either as limiting cases of lines intersecting o and making the angle θ zero or as limiting cases of lines ultraparallel to o and making the distance d zero. Since it follows from Equations 59 and 59a that $z\bar{z} = \tan^2\theta/2$ and $z\bar{z} = -\tanh^2 d/2$, respectively, we naturally associate with lines parallel to o, and directed in the same sense as o, **divisors of zero**, which are numbers of the form $u + ue$. In fact, lines parallel to o with positive or negative direction correspond to numbers $u + ev$, for which $v = u$ or $v = -u$, since it follows from Equation 59 and 59a that the relation $v = u$ is equivalent to the equation $s = \infty$ or $s' = \infty$ and the relation $v = -u$ is equivalent to the equation $s = -\infty$ or $s' = -\infty$. Further, it follows from the formulae of non-Euclidean trigonometry[54a] that the (oriented) distance $p = \{O, l\}$ from the pole O to a line l which intersects o (Figure 48), corresponding to the double

---

[54a] See H. S. M. Coxeter, *Non-Euclidean Geometry*, §12.8 (University of Toronto), 1957, or H. Schwerdtfeger, *Geometry of Complex Numbers*, §14 (University of Toronto, Oliver and Boyd), 1962.

number $z = u + ev = \tan \theta/2 \cdot (\cosh s + e \sinh s)$, is found from the following relation (compare with Equation 30):

$$\sinh p = \sinh s \cdot \sin \theta = \sinh s \frac{2 \tan \theta/2}{1 + \tan^2 \theta/2}$$

$$= \frac{2 \tan \theta/2 \cdot \sinh s}{1 + \tan^2 \theta/2} = \frac{2v}{1 + |z|^2} \tag{60}$$

So, with two lines n and n', parallel to o and separated from o by a distance $\{O, n\} = \{O, n'\} = p$, we must associate numbers $u + ev$ (where $v = \pm u$), for which $v = \frac{1}{2} \sinh p$, that is, numbers

$$z = \frac{\sinh p}{2}(1 + e) \qquad \text{and} \qquad z' = -\frac{\sinh p}{2}(1 - e)$$

Finally, starting from the relation $z_1 = -1/\bar{z}$, which connects double numbers z and z_1 corresponding to lines intersecting the axis o or ultraparallel to o and differing from each other only in direction, we associate with lines n_1 and n_1' **antiparallel** to o (lines parallel to o and in the opposite direction) and separated from o by a distance $\{O, n_1\} = \{O, n_1'\} = p$ the numbers

$$z_1 = \frac{2}{\sinh p_1} \omega_2 \qquad \text{and} \qquad z_1' = -\frac{2}{\sinh p_1} \omega_1$$

where ω_1 and ω_2 are numbers reciprocal to divisors of zero: $\omega_1 = 1/(1 + e)$ and $\omega_2 = 1/(1 - e)$. (We note that if n and n_1 are two lines which differ only in direction, then $p = \{O, n\} = -\{O, n_1\} = -p_1$.) With the polar axis o and the **antiaxis** o_1 (the line which differs from o only in direction) we associate the numbers 0 and ∞.

Up to now we have no lines corresponding to double numbers z such that $z\bar{z} = -1$ (since $\tanh^2 d/2 \neq 1$ and $\coth^2 d/2 \neq 1$ for any d). In order to extend the correspondence between lines of the Lobachevskii plane and double numbers to *all* the numbers, we introduce **lines at infinity** (infinite lines) of the Lobachevskii plane, which can be represented as tangents to the **absolute** Σ of

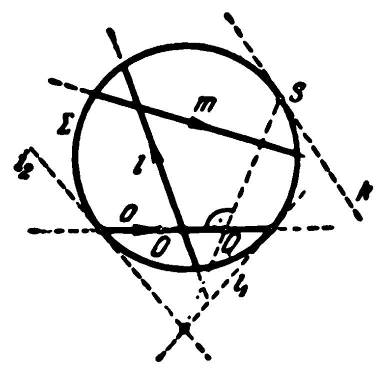


FIG. 49

the Klein model (Figure 49);[55] these lines, like the points at infinity in Section 11, have no orientation.[56] Such a line k "not parallel to o" (that is, different from the tangents to Σ at the points of intersection of Σ and o) is characterized by the fact that $d = \{k, o\} = \pm\infty$; here we must assume that $d = \infty$, if the point at infinity S of the Lobachevskii plane corresponding to k (the point of contact of k with the absolute Σ) lies to the right of o, and that $d = -\infty$ otherwise. The "common perpendicular" of k and o is naturally taken to be the line SQ, perpendicular to o; here the quantity $s' = \{O, Q\}$ can take any value, and so with each double number $z = \pm(\sinh s' + e \cosh s')$, such that $z\bar{z} = -1$, we can associate a definite *line at infinity* k. With lines at infinity i_1 and i_2 "parallel to o" (Figure 49) we associate the numbers $\sigma_1 = \omega_1/\omega_2$ and $\sigma_2 = \omega_2/\omega_1$.

[55] For the Klein model of the Lobachevskii plane see, for example, H. S. M. Coxeter, *Introduction to Geometry*, §16.2 (John Wiley, New York), 1961 and I. M. Yaglom, *Geometric Transformations*, Part III, Appendix to Chap. I (to be published by Random House, New York). This model of Lobachevskii geometry is often called the "Beltrami model" or the "Beltrami-Klein model," because the notable Italian geometer E. Beltrami (1835–1900) considered it earlier than that given by the famous German mathematician F. Klein (1849–1925); however, this part of Beltrami's work attracted little notice at the time, and only after the appearance of Klein's work was attention paid to it.

[56] This may be explained graphically by the fact that on these lines it is impossible to choose a segment whose direction would indicate the orientation of the line (there is nowhere to put the arrow!).

We have now set up a one-to-one correspondence between the set of (oriented and infinite) lines of the Lobachevskii plane and the set of double numbers (complemented by the numbers $c\omega_1$, $c\omega_2$, σ_1, σ_2, and ∞). The lines l, which intersect the polar axis o, correspond to double numbers $z = u + ev$ for which $z\bar{z} = u^2 - v^2 > 0$; that is, they correspond to numbers represented in the (u, v)-plane by points of the region denoted in Fig. 50 as region I. The lines m, ultraparallel to o and directed in

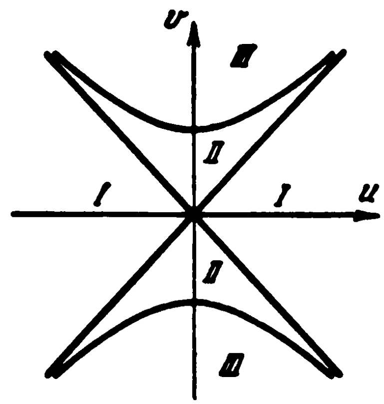


FIG. 50

the same sense as o from the common peprpendicular of o and m, correspond to numbers z for which $0 > z\bar{z} > -1$; these numbers are represented in the figure by points of the region II. Lines m_1, ultraparallel to o and directed in the sense opposite to o, from the common perpendicular of m_1 and o, correspond to numbers z for which $z\bar{z} < -1$; these are represented by points of the region III. Finally, lines n parallel to o correspond to numbers of zero modulus, which are represented by points of the lines $v = \pm u$, and lines n_1 antiparallel to o correspond to the numbers $c\omega_1$ and $c\omega_2$, which have no representation in the (u, v)-plane; lines at infinity k correspond to numbers z such that $z\bar{z} = -1$; that is, they correspond to numbers represented by points of the hyperbola $v^2 - u^2 = 1$ and to the two numbers σ_1 and σ_2.

It is obvious that, as in the case of a Euclidean plane, the relations (Equations 31)

$$z' = \bar{z}, \qquad z' = -z, \qquad z' = -\frac{1}{\bar{z}}$$

express the **symmetry about the point** O, the **symmetry about the line** o, and the **reorientation** (change of direction of all lines), respectively. An arbitrary **motion**, as we may show, is expressed here by the same formulae (see Equations 36a-d) as in the Euclidean case,

$$z' = \frac{Pz + Q}{-\bar{Q}z + \bar{P}}, \qquad \text{or} \quad z' = \frac{-Pz + Q}{\bar{Q}z + \bar{P}}$$

$$\text{or} \quad z' = \frac{P\bar{z} + Q}{-\bar{Q}\bar{z} + \bar{P}}, \qquad \text{or} \quad z' = \frac{-P\bar{z} + Q}{\bar{Q}\bar{z} + \bar{P}}$$

except that the variables z' and z and the coefficients P and Q are now not dual, but double, numbers; in this connection it is required, in addition, that the expression $P\bar{P} + Q\bar{Q}$ should be positive[57] (if P and Q are dual numbers, this condition is satisfied automatically, since the products $P\bar{P}$ and $Q\bar{Q}$ cannot be negative). Further, the (oriented) **angle** $\delta = \angle\{z_1, z_2\}$ between the lines z_1 and z_2 and the (oriented) **distance** $d = \{[z_1z_0], [z_2z_0]\}$ between the points of intersection of the lines z_1 and z_2 with the line z_0 are defined by the familiar formulae (Equations 37 and 38):[58]

$$\tan^2\frac{\delta}{2} = \frac{(z_2 - z_1)(\bar{z}_2 - \bar{z}_1)}{(1 + z_1\bar{z}_2)(1 + \bar{z}_1z_2)}$$

$$d = \arg\left(\frac{z_2 - z_0}{z_1 - z_0} : \frac{\bar{z}_0z_2 + 1}{\bar{z}_0z_1 + 1}\right)$$

---

[57] A transformation of the form $z' = (-Pz + Q)/(-\bar{Q}z + \bar{P})$, where $P\bar{P} + Q\bar{Q} > 0$, may be put into the form $z' = (Pz + Q)/(-\bar{Q}z + \bar{P})$, where $P\bar{P} + Q\bar{Q} < 0$, by means of the substitutions $P \to -eP$ and $Q \to eQ$, and similarly for the other two transformations.—TRANSL.

[58] If the lines z_1 and z_2 are ultraparallel, then the right-hand side of Equation 37 is negative, and this formula determines a complex value of the angle $\delta = i\Delta$ between these lines, where Δ is the shortest distance between z_1 and z_2 (cf. footnote 54); here we regard z_1 and z_2 as directed on one side of their common perpendicular. Similarly, if, for example, the line z_2 is ultraparallel to z_0, then the number in parentheses in Equation 38 will have the second of the forms given in Equation 39 of Chapter I, and the distance d, determined by Equation 38, will be complex: $d = D - i(\pi/2)$, where D is the distance from the point $[z_0z_1]$ to the projection of the common perpendicular of z_0 and z_2 on the line z_0 (see footnote 54).

It follows from Equation 38 that *the condition that the three lines* z_0, z_1, *and* z_2 *should meet in one point is that the ratio* $(z_0 - z_2)/(z_1 - z_2) : (\bar{z}_2 z_0 + 1)/(\bar{z}_2 z_1 + 1)$ *should be real.* Hence the *equation of a point* in the non-Euclidean geometry of Lobachevskii has the form (Equation 40):[59]

$$Az\bar{z} + Bz - \bar{B}\bar{z} - A = 0, \qquad A \text{ purely imaginary}$$

The following sets of (oriented and infinite) lines of the Lobachevskii plane should be called **cycles.**

(a), (b), (c) The set of lines which touch one of the (oriented) cycles considered in Section 11: **circles, horocycles,** and **equidistant curves.** Among the circles should be included *points* of the Lobachevskii plane (circles of radius 0); among the equidistant curves, *lines.*

(d) A **pencil of equal inclination,** which is a pencil consisting of all (oriented) lines that make a constant (oriented) angle with the fixed axis of the pencil. Among the pencils of equal inclination should be included *ultrainfinite points,* called *pencils of inclination* $\pi/2$ or *orthogonal pencils.*

(e) A **parallel pencil,** which is a pencil consisting of all lines parallel (in either direction) to the fixed axis of the pencil. Among the parallel pencils should be included *points at infinity* (parallel pencils whose axis is the line at infinity).

(f) The (nonoriented) **circle at infinity** (absolute) Σ.

If the term "cycle" is understood in this way, we obtain the result that the (necessary and sufficient) *condition that the four*

[59] The distance $d = \{[z_1 z_0], [z_2 z_0]\}$, determined by Equation 38, is equal to zero, not only when z_1 and z_2 meet z_0 in one point, but also when the lines z_0, z_1, and z_2 are ultraparallel and so have a common perpendicular (cf. footnote 58) or when all three lines are parallel to each other. Therefore, among the points of the non-Euclidean geometry of Lobachevskii should be included the points at infinity (points of the absolute of the Klein model), to which correspond pencils of parallel lines, and the ultrainfinite points (points lying outside the absolute of the Klein model), to which correspond pencils of ultraparallel lines.

(oriented) *lines* z_0, z_1, z_2, and z_3 *of the Lobachevskii plane should touch one cycle is that the cross-ratio* $W(z_0, z_1, z_2, z_3) = (z_0 - z_1)/(z_1 - z_2) : (z_0 - z_3)/(z_1 - z_3)$ *of these four lines should be real.* Hence it follows again that the *equation of every cycle can be written in the form* (Equation 14):

$$Az\bar{z} + Bz - \bar{B}\bar{z} + C = 0, \qquad A, C \text{ purely imaginary}$$

To decide whether the cycle, Equation 14, is a circle, a horocycle, an equidistant curve, a parallel pencil, or a pencil of equal inclination, we must find how many lines this cycle has in common with the circle at infinity (absolute) $z\bar{z} = -1$ (that is, the number of solutions of the system $z\bar{z} = -1$, $Bz - \bar{B}\bar{z} = A - C$) and whether the angle (Equation 37) between two neighboring lines of the cycle is real or imaginary. In this way we obtain the following results.

The cycle, Equation 14, *is a circle if*

$$AC + B\bar{B} > 0, \qquad (A - C)^2 - 4B\bar{B} < 0 \tag{61a}$$

It is a horocycle if

$$AC + B\bar{B} > 0, \qquad (A - C)^2 - 4B\bar{B} = 0, \qquad B \neq 0 \tag{61b}$$

It is an equidistant curve if

$$AC + B\bar{B} > 0, \qquad (A - C)^2 - 4B\bar{B} > 0 \tag{61c}$$

It is a parallel pencil if

$$AC + B\bar{B} = 0 \tag{61d}$$

It is a pencil of equal inclination if

$$AC + B\bar{B} < 0 \tag{61e}$$

It coincides with the absolute Σ *if*

$$A = C, \qquad B = 0 \tag{61f}$$

We have already seen that Equation 14 *represents an* (ordinary, infinite, or ultrainfinite) *point if* (Equation 43):

$$A + C = 0$$

These results can be used to prove many theorems in the non-Euclidean geometry of Lobachevskii. For example, exactly as on pp. 34–35 we can show that, *if* S_1, S_2, S_3, *and* S_4 *are four cycles of the Lobachevskii plane, and if* z_1 *and* w_1 *are the* (oriented) *common tangents of* S_1 *and* S_2, z_2 *and* w_2 *the* (oriented) *common tangents of* S_2 *and* S_3, z_3 *and* w_3 *the* (oriented) *common tangents of* S_3 *and* S_4, *and* z_4 *and* w_4 *the* (oriented) *common tangents of* S_4 *and* S_1, *and if* z_1, z_2, z_3, *and* z_4 *touch one cycle, then* w_1, w_2, w_3, *and* w_4 *also touch one cycle.*

The reader will be able to find for himself other applications of double numbers to the proof of theorems in the non-Euclidean geometry of Lobachevskii.

We note in conclusion that the representation of the set of (directed and infinite) lines of the Lobachevskii plane on the set of double numbers can be given in a somewhat different form. We return to what was said at the end of the preceding section about the Poincaré half-plane model of Lobachevskii geometry. In this model (nonoriented) points of the Lobachevskii plane are represented by points of one (the upper) half-plane, and lines are represented by semicircles with centers on the line o bounding the half-plane and by rays perpendicular to o; see Figure 51.

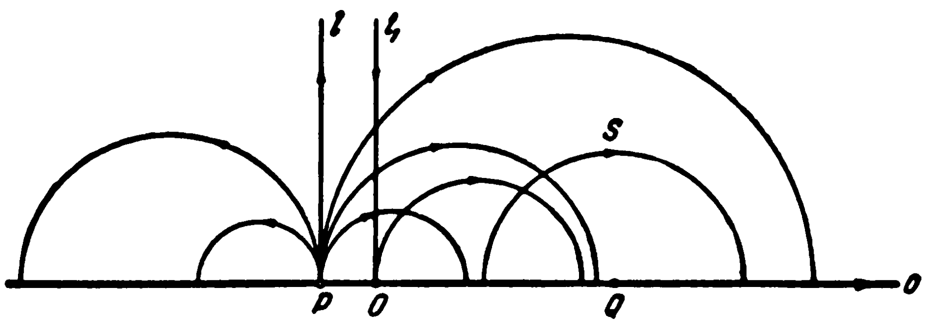


FIG. 51

Let us now agree to assign to the (oriented) line of the Lobachevskii plane, which is represented by a semicircle of radius r (where r can be positive or negative) with center at the point Q with abscissa x (that is, the point Q such that $\{O, Q\} = x$, where O is an origin of reference chosen on the oriented line o), the double number

$$z = x + er \tag{62}$$

Equation 62 sets up a correspondence between those (oriented) lines of the Lobachevskii plane which are represented by semicircles in the Poincaré model and double numbers $z = u + ev$, where $v \neq 0$. Here the *divisors of zero* $u \pm ue$ correspond to semicircles passing through the point O. Numbers $u + ev$ such that $u^2 - v^2 > 0$ (numbers which are represented by the first of the forms given in Equation 39 of Chapter I and which in Figure 50 are represented by points of the region I) corres-

pond to lines represented by semicircles for which the point O is exterior; numbers $u + ev$ such that $u^2 - v^2 < 0$ (numbers which are represented by the second of the forms given in Equation 39 of Chapter I and which in Figure 50 are represented by points of the regions II and III) correspond to lines represented by semicircles containing the point O inside them.[60] "Purely imaginary" numbers of the form ve correspond to lines represented by semicircles with center at the point O. Conjugate double numbers $z = u + ev$ and $\bar{z} = u - ev$ correspond to lines differing only in orientation (just as in the Poincaré half-plane model, discussed at the end of Section 11, conjugate complex numbers correspond to points differing only in orientation). It is clear also that, if we wish to extend our representation to lines at infinity of the Lobachevskii plane, represented by points of the absolute (cf. p. 120), then with a line, x, which plays the part of *tangent to the absolute o at the point Q* we naturally associate a real number $z = x$ (thus, to the *tangent to the absolute o at the point O* corresponds the number 0, and to the *tangent to the absolute o at its point at infinity* corresponds the number ∞).

So far we have not used the "singular" double numbers $c\omega_1$, $c\omega_2$, σ_1, and σ_2; on the other hand, lines represented on the Poincaré model by rays (and not by semicircles) have no corresponding numbers. However, it is clear that semicircles, y, which pass through the fixed point P and touch there the ray l, shown in Figure 51, correspond to double numbers of the form $w = (y - r) + re$. Since

$$\frac{1}{w} = \frac{(y - r) - re}{y^2 - 2yr} = \frac{-y/r + (1 + e)}{-y^2/r + 2y}$$

tends to the divisor of zero $(1/2y)(1 + e)$ as $| r |$ tends to ∞, we naturally associate with the ray l the number

$$z = 2y\omega_1 \tag{62a}$$

In exactly the same way it can be shown that with the ray l_1, opposite in direction to the ray l, we must associate the number $z_1 = 2y\omega_2$ where, as before, $z_1 = \bar{z}$ (see Equation 37a of Chapter I). The double numbers $\sigma_1 = (1 - e)/(1 + e)$ and $\sigma_2 = (1 + e)/(1 - e) = \bar{\sigma}_1$ are associated with the rays i (parallel to l) and i_1 (parallel to l_1) which pass through O.

We now find the angle $\delta = \angle\{z_1, z_2\}$ between two lines of the Lobachevskii plane which correspond to the double numbers

---

[60] Generally, the number $u^2 - v^2$ (the square of the modulus of the double number $z = u + ev$, taken with a suitable sign) is equal to the *power* of the semicircle corresponding to the number (in other words, the power of the point O with respect to the corresponding circle; see p. 44).

$z_1 = x_1 + er_1$ and $z_2 = x_2 + er_2$; these lines are represented on the Poincaré model by semicircles S_1 and S_2 having centers Q_1 and Q_2 (corresponding to x_1 and x_2) and radii r_1 and r_2 and intersecting at the point P. Depending on whether the orientations of the semicircles S_1 and S_2 are the same or different, we have

$$|\delta| = \angle Q_1PQ_2 \quad \text{or} \quad |\delta| = 180^\circ - \angle Q_1PQ_2$$

See Figure 52a, b; we note that the angle δ is equal to the *Euclidean* angle between the semicircles S_1 and S_2. But from the cosine theorem we have

$$\cos \angle Q_1PQ_2 = \frac{Q_1P^2 + Q_2P^2 - Q_1Q_2^2}{2Q_1P \cdot Q_2P} = \frac{r_1^2 + r_2^2 - (x_1 - x_2)^2}{2\,|r_1|\,|r_2|}$$

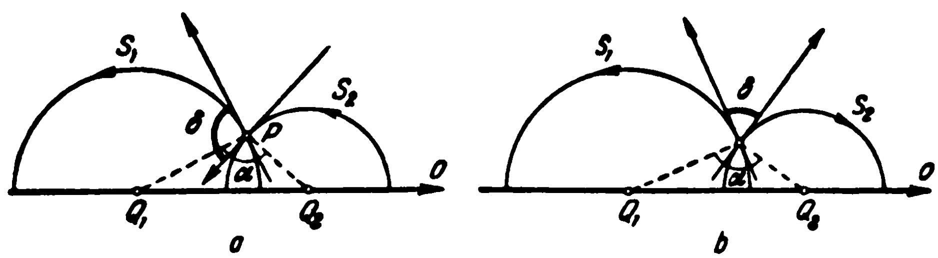


FIG. 52

Therefore, *in all cases*

$$\cos \delta = \frac{r_1^2 + r_2^2 - (x_1 - x_2)^2}{2r_1r_2}$$

whence we obtain (see Equation 56):

$$\tan^2 \frac{\delta}{2} = \frac{1 - \cos \delta}{1 + \cos \delta} = \frac{(x_1 - x_2)^2 - (r_1 - r_2)^2}{(r_1 + r_2)^2 - (x_1 - x_2)^2}$$

or

$$\tan^2 \frac{\delta}{2} = \frac{(z_2 - z_1)(\bar{z}_2 - \bar{z}_1)}{(\bar{z}_2 - z_1)(\bar{z}_1 - z_2)} \tag{63}$$

It is easy to verify that this equation remains valid when one of the lines is represented, not by a semicircle, but by a ray; in such case it is only necessary to substitute in Equation 63 for z_1, say, a number of the form $a/(1 + e)$ or $a/(1 - e)$.

We shall not write out the somewhat more complicated formula for the distance between two points (related to Equation 38). We merely note that here also it may be shown that *the necessary and sufficient condition that the four* (oriented) *lines* z_0, z_1, z_2, *and* z_3 *of the Lobachevskii plane should touch one cycle is that the cross-ratio* $W(z_0, z_1, z_2, z_3) = (z_0 - z_2)/(z_1 - z_2) : (z_0 - z_3)/(z_1 - z_3)$ *of these*

four lines should be real. Hence it follows that, for this representation of the set of (oriented and infinite) lines of the Lobachevskii plane on the set of double numbers, *the equations of cycles of the Lobachevskii plane*, as before, *have the form* (Equation 14):

$$Az\bar{z} + Bz - \bar{B}\bar{z} + C = 0, \qquad A, C \text{ purely imaginary}$$

To determine whether the cycle, Equation 14, is a circle, a horocycle, an equidistant curve, a parallel pencil, or a pencil of equal inclination, we must find the number of lines at infinity which touch the cycle and also whether the angle, Equation 63, between two neighboring lines of the cycle is real or imaginary. As a result, we arrive at the following theorem:

The cycle, Equation 14, *is a circle if* $AC + B\bar{B} > 0$ *and* $(B - \bar{B})^2 - 4AC < 0$, *a horocycle if* $AC + B\bar{B} > 0$, $(B - \bar{B})^2 - 4AC = 0$, *and* $A^2 + C^2 \neq 0$, *an equidistant curve if* $AC + B\bar{B} > 0$ *and* $(B - \bar{B})^2 - 4AC > 0$, *a parallel pencil if* $AC + B\bar{B} = 0$, *a pencil of equal inclination if* $AC + B\bar{B} < 0$, *and the absolute o if* $A = C = 0$ and $B - \bar{B} < 0$.

It is not difficult to verify also that *the cycle* (Equation 14) *is a point* (ordinary, infinite, or ultrainfinite) *if and only if the following* (Equation 57) *is satisfied*:

$$B + \bar{B} = 0$$

Finally, we find that *the motions of the Lobachevskii plane in the complex* (more precisely, double) *line coordinates considered here can be written as*

$$z' = \frac{az + b}{cz + d} \tag{64}$$

where a, b, c, and d are real numbers and $ad - bc \neq 0$ (see Equation 58).

The connection between the two representations, considered in this section, of the set of lines of the Lobachevskii plane on the set of double numbers will be established in Section 18.

CHAPTER III

Circular Transformations and Circular Geometry

§13. Ordinary Circular Transformations (Möbius Transformations)

In this section we shall consider arbitrary linear-fractional functions of a complex variable z and *linear-fractional transformations* of the plane, which correspond to them by virtue of what was said in Section 7; these are the transformations expressed by the formulae[61]

$$z' = \frac{az + b}{cz + d} \tag{1}$$

$$z' = \frac{a\bar{z} + b}{c\bar{z} + d} \tag{1a}$$

If $\Delta = \left| \begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right| = 0$, so that $a = kc$ and $b = kd$, the functions defined by Equations 1 and 1a reduce to $z' = k$; thus, we are only interested in the case in which $\Delta = \left| \begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right| \neq 0$, and henceforth we shall always assume that the latter condition is satisfied. In this case the transformations, Equations 1 and 1a, will be one-to-one transformations of the plane of the complex variable, extended by introducing the number $1/0 = \infty$; in fact, to each number z corresponds a unique number z', determined by Equation 1 or 1a,

[61] See, for example, H. S. M. Coxeter, *Introduction to Geometry*, §9.7 (John Wiley, New York), 1961, and H. Schwerdtfeger, *Geometry of Complex Numbers*, §6 (University of Toronto, Oliver and Boyd), 1962.

and to each value of z' corresponds a unique value of z, found from the same equations:

$$z = \frac{dz' - b}{-cz' + a} \qquad \text{or} \qquad z = \frac{\bar{d}\bar{z}' - \bar{b}}{-\bar{c}\bar{z}' + \bar{a}} \tag{2}$$

In particular, with the number $z = \infty$ the Equations 1 and 1a associate the value[62] $z' = a/c$ and, by virtue of the same formulae, to the number $z = -d/c$ or $z = -\bar{d}/\bar{c}$, such that $cz + d = 0$ or $\bar{c}z + \bar{d} = 0$, corresponds the value $z' = \infty$.

In a special case the linear-fractional transformations are **linear transformations**

$$z' = pz + q, \quad p \neq 0 \qquad \text{or} \qquad z' = p\bar{z} + q, \quad p \neq 0 \tag{3}$$

which we obtain by substituting $c = 0$ into Equations 1 and 1a and putting $p = a/d$ and $q = b/d$. Sometimes Equation 1 is called a **direct** linear-fractional transformation or **homography** and Equation 1a is called an **opposite** linear-fractional transformation or **antihomography**.

We have seen that geometrically a linear transformation represents a **similarity**, which is a combination of a *spiral similarity* (a dilatation and rotation with common center O) and a *translation* and, possibly, a symmetry about a line; in particular, if $|p| = 1$, a linear transformation is a *motion* (see the beginning of Section 7). We have also considered some of the simplest concrete examples of linear transformations, namely the transformations

$$z' = -z \qquad \text{and} \qquad z' = \bar{z} \tag{4a,b}$$

which consist of the *symmetry about the point* O and the *symmetry about the line* o, and also the transformations

$$z' = z + q \qquad \text{and} \qquad z' = pz \tag{5a,b}$$

which are a *translation* and a *spiral similarity* (see the beginning of Section 7). Here we shall study in greater detail

[62] See footnote 4.

the geometrical properties of arbitrary linear-fractional transformations.

We note first of all that the **product** (the result of successive applications) *of two linear-fractional transformations is again a linear-fractional transformation*; *the* **identical** (*or* **unit**) *transformation*, which leaves all points unaltered, *is a special case of a linear-fractional transformation*; *the* **inverse** *of a linear-fractional transformation* (the one which takes each point z' of the plane to the point z, from which z' was obtained as a result of the original transformation) *is again linear-fractional.* In fact, for example, if

$$z_1 = \frac{az + b}{cz + d} \quad \text{and} \quad z' = \frac{a_1 z_1 + b_1}{c_1 z_1 + d_1} \tag{6}$$

then

$$z' = \frac{a_1(az + b)/(cz + d) + b_1}{c_1(az + b)/(cz + d) + d_1} = \frac{(a_1 a + b_1 c)z + (a_1 b + b_1 d)}{(c_1 a + d_1 c)z + (c_1 b + d_1 d)} \tag{6a}$$

and similarly for the product of a homography and an antihomography or the product of two antihomographies.[63] The identical transformation is expressed by the formula

$$z' = z \tag{7}$$

which is a special case of Equation 1 (when $b = c = 0$ and $a = d = 1$). Finally, the transformations inverse to Equations 1 and 1a have the following forms (compare with Equation 2):

$$z' = \frac{dz - b}{-cz + a} \quad \text{or} \quad z' = \frac{\bar{d}\bar{z} - \bar{b}}{-\bar{c}\bar{z} + \bar{a}} \tag{8}$$

We single out now the following fundamental property of linear-fractional transformations: *if z_1', z_2', z_3', and z_4' are the four points of the plane which correspond to four given points*

[63] It is not difficult to verify that *the product of two homographies or two antihomographies is a homography;* on the other hand, *the product of a homography and an antihomography* (in either order) *is always an antihomography.*

z_1, z_2, z_3, *and* z_4 *under the transformation* (Equation 1 or 1a), *then*

$$W(z_1', z_2', z_3', z_4') = W(z_1, z_2, z_3, z_4) \tag{9}$$

or

$$W(z_1', z_2', z_3', z_4') = \overline{W(z_1, z_2, z_3, z_4)} \tag{9a}$$

where $W(z_1, z_2, z_3, z_4) = (z_1 - z_3)/(z_2 - z_3) : (z_1 - z_4)/(z_2 - z_4)$ *is the cross-ratio of the four points* (the **property of the invariance of the cross-ratio**). In fact, from Equation 1, for example, we obtain

$$\begin{aligned}
W(z_1', z_2', z_3', z_4') &= \frac{z_1' - z_3'}{z_2' - z_3'} : \frac{z_1' - z_4'}{z_2' - z_4'} \\
&= \frac{(az_1 + b)/(cz_1 + d) - (az_3 + b)/(cz_3 + d)}{(az_2 + b)/(cz_2 + d) - (az_3 + b)/(cz_3 + d)} \\
&\quad : \frac{(az_1 + b)/(cz_1 + d) - (az_4 + b)/(cz_4 + d)}{(az_2 + b)/(cz_2 + d) - (az_4 + b)/(cz_4 + d)} \\
&= \frac{[(ad - bc)(z_1 - z_3)]/(cz_1 + d)}{[(ad - bc)(z_2 - z_3)]/(cz_2 + d)} \\
&\quad : \frac{[(ad - bc)(z_1 - z_4)]/(cz_1 + d)}{[(ad - bc)(z_2 - z_4)]/(cz_2 + d)} \\
&= \frac{z_1 - z_3}{z_2 - z_3} : \frac{z_1 - z_4}{z_2 - z_4} = W(z_1, z_2, z_3, z_4)
\end{aligned}$$

and similarly it may be verified that $W(z_1', z_2', z_3', z_4') = \overline{W(z_1, z_2, z_3, z_4)}$ if z and z' are connected by Equation 1a.

From the property of the invariance of the cross-ratio it follows immediately that *a linear-fractional transformation takes four points, lying on one circle or line, into four points, also lying on one circle or line* (the **circular property of linear-fractional transformations**). In fact, since the cross-ratio $W(z_1, z_2, z_3, z_4)$ of the original points is real, it follows that the cross-ratio $W(z_1', z_2', z_3', z_4')$ of the transformed points also is real, from which our statement follows (see the beginning of Section 7). In turn, it follows from this that *a linear-fractional transformation*

takes each circle or line of the plane into a circle or line.[64] This is the reason why linear-fractional transformations of the plane are also called **circular transformations** (we may also speak of *direct* and *opposite* circular transformations[65]) and, since they were first studied thoroughly by the German geometer A. F. Möbius (1790–1868), they are often called **Möbius circular transformations**.

We may show that *there exists a unique homography* (Equation 1), *which takes three given points* z_1, z_2, *and* z_3 *into three other given points* w_1, w_2, *and* w_3. In fact, if this transformation takes the the points z_1, z_2, and z_3 into the points w_1, w_2, and w_3 and takes an arbitrary point z of the plane into a point z', then, by virtue of what was shown above,

$$W(z, z_1, z_2, z_3) = W(z', w_1, w_2, w_3)$$

or

$$\frac{z' - w_2}{w_1 - w_2} : \frac{z' - w_3}{w_1 - w_3} = \frac{z - z_2}{z_1 - z_2} : \frac{z - z_3}{z_1 - z_3} \tag{10}$$

But Equation 10 defines a linear-fractional transformation: if z' is expressed in terms of z, we obtain

$$z' = \frac{Az + B}{Cz + D}$$

where

$$\begin{aligned}
A &= w_1w_2(z_1 - z_2) + w_2w_3(z_2 - z_3) + w_3w_1(z_3 - z_1)\\
B &= w_1w_2z_3(z_2 - z_1) + w_2w_3z_1(z_3 - z_2) + w_3w_1z_2(z_1 - z_3)\\
C &= (w_2z_1 - w_1z_2) + (w_3z_2 - w_2z_3) + (w_1z_3 - w_3z_1)\\
D &= z_1z_2(w_1 - w_2) + z_2z_3(w_2 - w_3) + z_3z_1(w_3 - w_1)
\end{aligned}$$

---

[64] It is not difficult to show by calculation that, for example, the circular transformation, Equation 1, takes the circle or line whose equation has the form $Az\bar{z} + Bz - \bar{B}\bar{z} + C = 0$, where A and C are purely imaginary (Equation 14, Chapter II), into the circle or line $A'z\bar{z} + B'z - \bar{B}'\bar{z} + C' = 0$, where $A' = Aa\bar{a} + Ba\bar{c} - \bar{B}\bar{a}c + Cc\bar{c}$, $B' = Aa\bar{b} + Ba\bar{d} - \bar{B}c\bar{b} + Cc\bar{d}$, and $C' = Ab\bar{b} + Bb\bar{d} - \bar{B}\bar{b}d + Cd\bar{d}$.

[65] It may be shown that *all* transformations of the plane of the complex variable which take circles or lines into circles or lines are exhausted by the linear-fractional transformations, Equations 1 and 1a; see H. S. M. Coxeter, *Introduction to Geometry*, §9.7 (John Wiley, New York), 1961.

In exactly the same way it may be shown that *there exists a unique antihomography* (Equation 1a), *which takes* z_1, z_2, *and* z_3 *into* w_1, w_2, *and* w_3 ; this transformation is given by the formula

$$W(z', w_1, w_2, w_3) = \overline{W(z, z_1, z_2, z_3)}$$

or

$$\frac{z' - w_2}{w_1 - w_2} : \frac{z' - w_3}{w_1 - w_3} = \frac{\bar{z} - \bar{z}_2}{\bar{z}_1 - \bar{z}_2} : \frac{\bar{z} - \bar{z}_3}{\bar{z}_1 - \bar{z}_3} \tag{10a}$$

This reasoning shows that *the necessary and sufficient condition that four given points* z_1, z_2, z_3, *and* z_4 *may be taken into four other points* w_1, w_2, w_3, *and* w_4 *by a circular transformation is*

$$W(w_1, w_2, w_3, w_4) = W(z_1, z_2, z_3, z_4)$$

or

$$W(w_1, w_2, w_3, w_4) = \overline{W(z_1, z_2, z_3, z_4)}$$

It follows from this that *any circle or line may* (in many ways) *be taken by a circular transformation* (Equation 1 or 1a) *into any other given circle or line*; one must merely ensure that any three points of the first circle go into three (arbitrary) points of the second circle. In particular, *any circle can be taken into a line in infinitely many ways*; this is a result which is often useful. Thus, from the point of view of circular transformations, all circles and lines are completely equivalent; therefore, in problems connected with circular transformations we usually do not distinguish between lines and circles, regarding lines as special cases of circles, "circles with infinite radius." In future we shall often simply speak of a **circle** when it would be more correct to say "circle or line."

We shall now explain the geometrical meaning of the cross-ratio $W(z_1, z_2, z_3, z_4)$ of four points z_1, z_2, z_3, and z_4. We know already that, if this cross-ratio is real (that is, if the argument arg W of the cross-ratio W is equal to zero), then the points

z_1, z_2, z_3, and z_4 lie on one circle (or line). In the general case we have, obviously,

$$\begin{aligned}\arg W(z_1, z_2, z_3, z_4) &= \arg\left(\frac{z_1 - z_3}{z_2 - z_3} : \frac{z_1 - z_4}{z_2 - z_4}\right) \\ &= \arg\frac{z_1 - z_3}{z_2 - z_3} - \arg\frac{z_1 - z_4}{z_2 - z_4} \\ &= \angle\{[z_3z_2], [z_3z_1]\} - \angle\{[z_4z_2], [z_4z_1]\}\end{aligned}$$

See Equations 8, Chapter II. Let us consider two circles (one or both of them may be a line) passing through the points z_1, z_2, z_3 and z_1, z_2, z_4. These circles (see Figure 53) will be denoted by

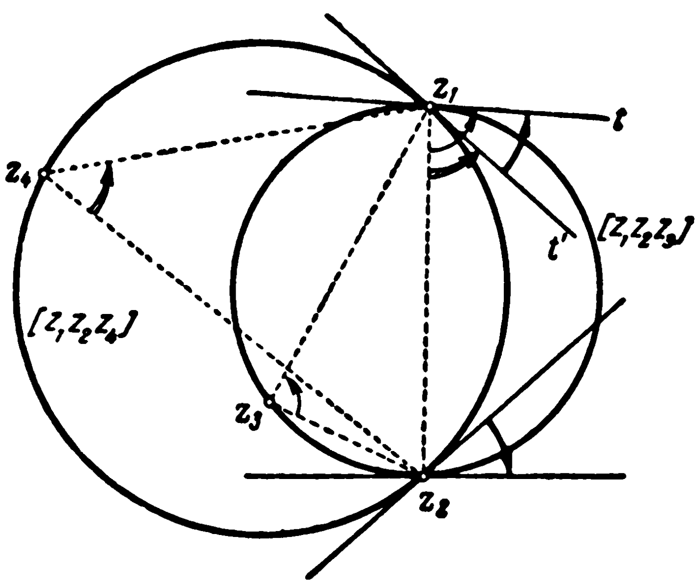


FIG. 53

$[z_1z_2z_3]$ and $[z_1z_2z_4]$; in general, the circle passing through the points z, v, and w will be called the circle $[zvw]$. We draw the tangents t and t' to these circles at the point z_1. From a well-known theorem about angles in a circle it follows, independently of the positions of the points z_3 and z_4 on the circles $[z_1z_2z_3]$ and $[z_1z_2z_4]$, that

$$\angle\{[z_3z_2], [z_3z_1]\} = \angle\{[z_1z_2], t\}$$

and

$$\angle\{[z_4z_2], [z_4z_1]\} = \angle\{[z_1z_2], t'\}$$

We have, therefore,

$$\arg W(z_1, z_2, z_3, z_4) = \angle\{[z_1z_2], t\} - \angle\{[z_1z_2], t'\}$$
$$= \angle\{t', t\} \tag{11}$$

The angle between the tangents to the circles S_1 and S_2 at a point of intersection z is called the **angle between the circles** S_1 and S_2 and is denoted by $\angle(S_1, S_2)$; if the oriented angle between the tangents is considered, we speak of the **oriented angle** $\angle\{S_1zS_2\}$ **between the circles** S_1 and S_2. Thus we see that *the argument* $\arg W(z_1, z_2, z_3, z_4)$ *of the cross-ratio* $W(z_1, z_2, z_3, z_4)$ *of four points* z_1, z_2, z_3, *and* z_4 *is equal to the* (oriented) *angle* $\angle\{[z_1z_2z_4]z_1[z_1z_2z_3]\}$ *between the circles* $[z_1z_2z_4]$ *and* $[z_1z_2z_3]$.

From the property of the invariance of the cross-ratio of four points it follows that *a direct circular transformation does not alter the oriented angle between intersecting circles, and an opposite circular transformation alters the sign* (the direction of rotation), *but not the absolute value, of this angle.*[66] This important property of circular transformations is often formulated briefly in the following way: *Angles between circles are preserved by circular transformations*; see Figure 54.[67] In particular, *intersecting circles*

[66] In other words, if a direct or opposite circular transformation takes the circles S_1 and S_2 into the circles S_1' and S_2' and a point of intersection z_1 of S_1 and S_2 into a point z_1', then $\angle\{S_1z_1S_2\} = \angle\{S_1'z_1'S_2'\}$ or $\angle\{S_1z_1S_2\} = -\angle\{S_1'z_1'S_2'\}$, respectively. We may note that if two circles S_1 and S_2 intersect at points z_1 and z_2, then $\angle\{S_1z_1S_2\} = -\angle\{S_1z_2S_2\}$; see Figure 53. Therefore, when referring to the oriented angle between two intersecting circles it is necessary to indicate the point of intersection at which the angle is being considered (an unoriented angle between circles does not depend on the choice of point of intersection).

[67] Since the angle between two *arbitrary curves* γ_1 and γ_2 which intersect at a point z (by definition this angle coincides with the angle between the tangents to γ_1 and γ_2 at the point z) is equal to the angle between the circles S_1 and S_2 which touch these curves at z (Figure 54), and since a circular transformation which takes the curves γ_1 and γ_2 into new curves γ_1' and γ_2' takes S_1 and S_2 into circles S_1' and S_2' which touch γ_1' and γ_2', it follows that *angles between arbitrary curves are preserved*

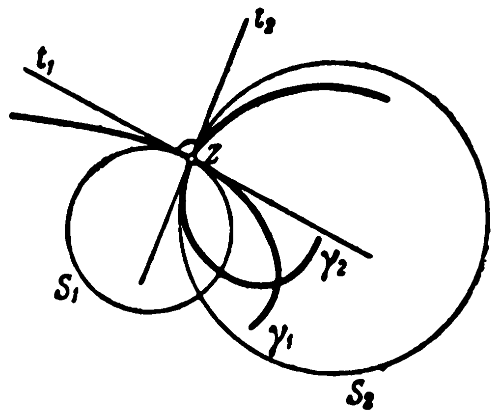


FIG. 54

(circles with a nonzero angle between them) *go into intersecting circles under a circular transformation.*

We now turn to the modulus $|W(z_1, z_2, z_3, z_4)|$ of the cross-ratio W of four points of the plane. By virtue of Equation 6, Chapter II, we have

$$\begin{aligned} |W(z_1, z_2, z_3, z_4)| &= \left| \frac{z_1 - z_3}{z_2 - z_3} : \frac{z_1 - z_4}{z_2 - z_4} \right| \\ &= \frac{|z_1 - z_3|}{|z_2 - z_3|} : \frac{|z_1 - z_4|}{|z_2 - z_4|} \\ &= \frac{(z_1, z_3)}{(z_2, z_3)} : \frac{(z_1, z_4)}{(z_2, z_4)} \qquad (12) \end{aligned}$$

where (z_1, z_3), (z_2, z_3), and so on, are the distances between the points z_1 and z_3, and z_2 and z_3, and so on. The real number $(z_1, z_3)/(z_2, z_3) : (z_1, z_4)/(z_2, z_4)$ will be called the **cross-ratio of the distances between the four points** z_1, z_2, z_3, and z_4 and will be denoted by $\tilde{W}(z_1, z_2, z_3, z_4)$. Thus we have

$$|W(z_1, z_2, z_3, z_4)| = \tilde{W}(z_1, z_2, z_3, z_4) \qquad (12a)$$

or, in words, *the modulus* $|W(z_1, z_2, z_3, z_4)|$ *of the cross-ratio of the four points* z_1, z_2, z_3, *and* z_4 *is equal to the cross-ratio of the*

---

by circular transformations. All transformations which have this property are called **conformal transformations**. Thus *circular transformations are conformal transformations.*

distances between these points. From the property of the invariance of the cross-ratio we may conclude that circular transformations preserve the cross-ratios of the distances between tetrads of points.

We may now reformulate the necessary and sufficient condition that four points z_1, z_2, z_3, and z_4 may be taken by a circular transformation into four points w_1, w_2, w_3, and w_4; namely, *the angle between the circles* $[z_1z_2z_3]$ *and* $[z_1z_2z_4]$ *should be equal to the angle between the circles* $[w_1w_2w_3]$ *and* $[w_1w_2w_4]$, and *the cross-ratio of the distances between the points* z_1, z_2, z_3, *and* z_4 *should be equal to the cross-ratio of the distances between the points* w_1, w_2, w_3, *and* w_4.

We shall now consider the question of the *geometrical description* of arbitrary circular transformations. As we know, a linear transformation (Equations 3) consists of a *similarity.* The product or two or more linear transformations is again a linear transformation;[68] therefore it is impossible to reduce all circular transformations to similarities only.

The simplest circular transformations, apart from similarities, are

$$z' = \frac{1}{z} \qquad \text{and} \qquad z' = \frac{1}{\bar{z}} \tag{13a,b}$$

which may also be expressed by the equations

$$\begin{aligned} \arg z' &= -\arg z, & |z'| &= \frac{1}{|z|} \\ \arg z' &= \arg z, & |z'| &= \frac{1}{|z|} \end{aligned} \tag{14a,b}$$

Of these two transformations the one with the simpler geometrical meaning is Equation 13b or 14b, called the **unit inversion.** Under this inversion each point z of the plane goes into a point z' of the ray Oz (the fact that z' lies on the ray Oz follows from the equation $\arg z' = \arg z$), such that

$$(O, z') = 1/(O, z) \qquad \text{or} \qquad (O, z) \cdot (O, z') = 1 \tag{15}$$

See Figure 55; the inverse of the point O is the "point" ∞.

---

[68] So, for example, if $z_1 = az + b$ and $z' = a_1z_1 + b_1$, then $z' = Az + B$, where $A = a_1a$ and $B = a_1b + b_1$.

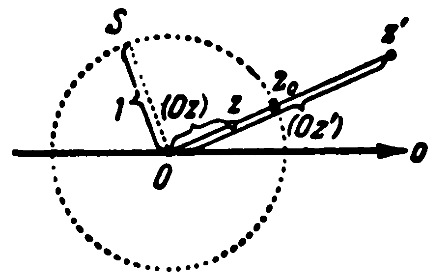


FIG. 55

Inversion is one of the simplest circular transformations. Just as any circular transformation does, it takes each circle or line into a circle or line; inversion preserves angles between (intersecting) circles (moreover, the signs of oriented angles are altered by inversion, inversion being an opposite circular transformation); the cross-ratio of the distances of four points of the plane is unaltered by inversion. Each point of the unit circle $z\bar{z} = 1$ goes into itself under the inversion. This property and also the fact that each point z exterior to the circle $z\bar{z} = 1$ goes into an interior point z', which lies on the ray Oz (see Figure 55) and is such that the radius $(O, z_0) = 1$ of the unit circle is a mean proportional between the lengths of the segments (O, z) and (O, z'), and, conversely, the fact that the point z' goes into the point z, give the reason why the transformation (Equation 13b or 14b) is also called the **symmetry with respect to the unit circle**; see Figure 56.[69]

The transformation given by Equation 13a does not merit special consideration; it is obviously the product of the unit inversion (symmetry about the unit circle, Equation 13b) and the symmetry about the axis o (Equation 4b). A transformation which also reduces to those considered earlier is an **inversion of arbitrary power k**,

$$z' = \frac{k}{\bar{z}}, \qquad k \text{ real} \tag{16}$$

[69] A further reason is the fact that *any circle which passes through two points z and z' which correspond to each other under the unit inversion is orthogonal to the unit circle $z\bar{z} = 1$* (Figure 56a), just as any circle which passes through two points z and z' symmetrical about a line l is orthogonal to l (Figure 56b).

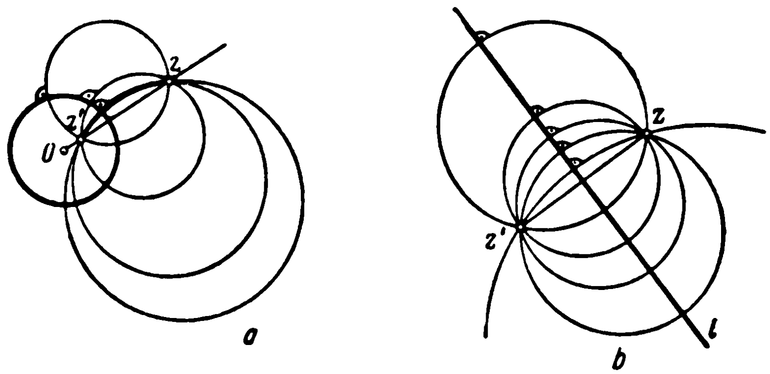


FIG. 56

which is the product of the unit inversion (Equation 13b) and the dilatation $z' = kz$ with ratio k.[70] In general, *each circular transformation is the product of a similarity and the* (unit) *inversion* [¶D]; the proof of this proposition will be our main aim.

First of all we define more precisely the formulae which describe inversions. Geometrically a unit inversion is defined as a transformation which takes an arbitrary point z of the plane into the point z' of the ray Oz (where O is a fixed point) such that Equation 15 is satisfied. Thus, in the geometrical description of inversion the point O plays an important role; this point is called the **center of the inversion**. If the center of the inversion coincides with the origin of the coordinate system, then the inversion is expressed by Equation 13b; if an arbitrary point w of the plane is the center of the inversion (Figure 57), then the inversion is, obviously, expressed as

$$z' - w = \frac{1}{\overline{z - w}} \qquad \text{or} \qquad z' = \frac{w\bar{z} + (1 - w\bar{w})}{\bar{z} - \bar{w}} \tag{17}$$

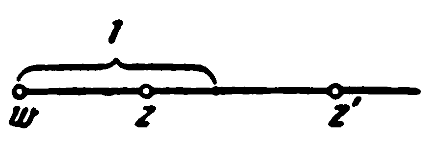


FIG. 57

[70] If $k > 0$ the inversion of power k is also called the **symmetry with respect to the circle of radius** $\sqrt{k}$ (the circle $z\bar{z} = k$) [¶E]. An inversion of power $k < 0$ can be regarded as the product of an inversion of positive power $|k|$ and the symmetry about the origin O (Equation 4a).

It is not difficult to verify that *every direct circular transformation* $z' = (az + b)/(cz + d)$ *other than a similarity, where* $c \neq 0$, *is the product of a similarity* $z' = p\bar{z} + q$, *where*

$$p = -\frac{\bar{c}^2}{\bar{\Delta}}, \qquad q = \frac{a\bar{a}\bar{d} - c\bar{c}\bar{d} - a\bar{c}\bar{b}}{c\bar{\Delta}}, \qquad \Delta = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$$

and an inversion (Equation 17), *where*

$$w = \frac{a}{c}$$

and that *every opposite circular transformation*

$$z' = \frac{a\bar{z} + b}{c\bar{z} + d}$$

where $c \neq 0$, *is the product of a similarity* $z' = pz + q$, *where*

$$p = \frac{c^2}{\Delta}, \qquad q = \frac{a\bar{a}d - c\bar{c}d - \bar{a}cb}{\bar{c}\Delta}$$

and the same inversion (Equation 17). In fact, if, for example,

$$z_1 = -\frac{\bar{c}^2}{\bar{\Delta}}\bar{z} + \frac{a\bar{a}\bar{d} - c\bar{c}\bar{d} - a\bar{c}\bar{b}}{c\bar{\Delta}}$$

is a similarity which takes the point z into the point z_1, and if

$$z' = \frac{(a/c)\bar{z}_1 + (1 - a\bar{a}/c\bar{c})}{\bar{z}_1 - \bar{a}/\bar{c}}$$

is an inversion which takes the point z_1 into the point z', then we have

$$\begin{aligned} z' &= \frac{(a/c)[-(c^2/\Delta)z + (a\bar{a}d - c\bar{c}d - \bar{a}cb)/\bar{c}\Delta] + (1 - a\bar{a}/c\bar{c})}{-(c^2/\Delta)z + (a\bar{a}d - c\bar{c}d - \bar{a}cb)/\bar{c}\Delta - \bar{a}/\bar{c}} \\ &= \frac{az + b}{cz + d} \end{aligned}$$

Thus, all circular transformations other than similarities reduce to the unit inversion.[71] In exactly the same way, we may prove the result about opposite circular transformations.

At the end of this section we shall dwell on some of the main features of circular transformations. We have seen that the set of all circular transformations:

(a) Contains the product of any two transformations of the set.

(b) Contains the transformation inverse to any transformation of the set.

(c) Contains the identical (unit) transformation.

A set of transformations with all these properties is called a **group** of transformations. Thus, *circular transformations form a group.*

From the geometrical properties of plane figures let us select those which are preserved by circular transformations; among these properties are, say, the property of a curve's being a *circle or line* (but not the property of a curve's being a line, since a line can go into a circle under a circular transformation) and the property of two circles of intersecting at a definite angle α. These properties may be called **circular properties of figures**, and the study of these properties may be called **circular geometry**.

The definition of circular geometry is related to the definition of ordinary geometry as the study of those properties of figures which do not depend on the position of the figure in the plane: that is, those which are *preserved by all motions* (we note, incidentally, that the set of all motions obviously forms a group).[72] Because the set of circular transformations is *wider* than the set

---

[71] From what has been proved it follows that *every circular transformation other than a similarity can be expressed as the product of an inversion of some power k and a motion.*

[72] In this connection, see H. S. M. Coxeter, *Introduction to Geometry*, Chap. 5, first paragraph (John Wiley, New York), 1961, and I. M. Yaglom, *Geometric Transformations*, introductions to Parts I, II, III (Random House, New York).

of motions (since every motion $z' = pz + q$, where $p\bar{p} = 1$, is a special case of a circular transformation, while a circular transformation need not be a motion), circular geometry is a part of all geometry. The last statement should be understood to mean that the circular properties of figures are only *some* of the properties which are studied in ordinary (Euclidean) geometry.

From the point of view of circular geometry, figures in the plane which can be transformed into each other by circular transformations are indistinguishable (they have the same properties); in circular geometry such figures may be called *identical* or, as we shall say, **equal**. Since circular transformations form a group, in circular geometry the following obtain:

(a) If the figure Φ is equal to the figure Φ_1, and the figure Φ_1 is equal to the figure Φ_2, then the figure Φ is equal to the figure Φ_2 (since Φ is taken into Φ_2 by the product of the transformations which take Φ into Φ_1 and Φ_1 into Φ_2).

(b) If the figure Φ is equal to the figure Φ_1, then the figure Φ_1 is equal to the figure Φ (since Φ_1 is taken into Φ by the inverse of the transformation which takes Φ into Φ_1).

(c) Every figure Φ is equal to itself (since Φ is taken into itself by the identical transformation, which is a circular transformation).

Thus, equality of figures in circular geometry has these three properties, and only if they are satisfied may the term "equality" be used. The fact that circular transformations form a group is important, because these three properties, defined by means of circular transformations, follow from it.

The concept of circular geometry selects a certain class of geometrical properties of figures, which can be studied by the same methods. In particular, in proving circular properties of figures it is very helpful to use circular transformations, by means of which we can sometimes considerably simplify the corresponding diagram; in fact, from the point of view of circular geometry, all diagrams obtainable from each other by means of circular transformations are identical, and so we may use

whichever we please. The reader will meet a number of examples of the use of circular transformations in the next section.

There is another approach to circular geometry which, in the proof of the circular properties of figures, allows only the use of those theorems and concepts that relate to circular geometry (that preserve their meaning under circular transformations). From this point of view, none of the diagrams obtainable from each other by circular transformations can have any preference over the others, since the figures represented by these diagrams have exactly the same circular properties [¶F]. The restriction in the number of properties which we are allowed to use in the construction of circular geometry naturally complicates the problem of proving theorems in this geometry; this is partially compensated for by the fact that the limitation on the number of possible proofs can sometimes facilitate the discovery of the proper method. The value of this approach to circular geometry is that it enables us to consider this geometry an independent subject, just as we do ordinary (Euclidean) geometry.

This subject represents a new branch of geometry like, say, the non-Euclidean geometry of Lobachevskii.

*§14. Applications and Examples

We have seen that any triangle $\overline{z_1 z_2 z_3}$ can be taken, by a suitably chosen circular transformation, into any other triangle $\overline{w_1 w_2 w_3}$, in the sense that the points z_1, z_2, and z_3 can be taken into the points w_1, w_2, and w_3 (it is to be understood that, as a rule, the sides of the first triangle do not go into the sides of the second, but into circles). This does not hold for quadrangles: *in order that the quadrangle* $\overline{z_1 z_2 z_3 z_4}$ *may be taken by a circular transformation into the quadrangle* $\overline{w_1 w_2 w_3 w_4}$ (the vertices of the first into the vertices of the second), *it is necessary* (*and sufficient*) *that*

$$W(w_1, w_2, w_3, w_4) = W(z_1, z_2, z_3, z_4)$$

or that

$$W(w_1, w_2, w_3, w_4) = \overline{W(z_1, z_2, z_3, z_4)}$$

or, to put it another way, that

$$\angle([w_1w_2w_3]w_1[w_1w_2w_4]) = \angle([z_1z_2z_3]z_1[z_1z_2z_4])$$

and

$$\tilde{W}(w_1, w_2, w_3, w_4) = \tilde{W}(z_1, z_2, z_3, z_4)$$

See p. 135 ff.; the angles are not oriented, since the direction of angles is altered by opposite circular transformations.

But (see pp. 135–137):

$$\angle\{[z_1z_2z_3]z_1[z_1z_2z_4]\} = \angle\{[z_3z_2], [z_3z_1]\} - \angle\{[z_4z_2], [z_4z_1]\}$$

and

$$\tilde{W}(z_1, z_2, z_3, z_4) = \frac{(z_1, z_3)(z_2, z_4)}{(z_2, z_3)(z_1, z_4)}$$

Moreover, if, for simplicity of drawing, we restrict ourselves to the case of convex quadrangles $\overline{z_1z_3z_2z_4}$ and $\overline{w_1w_3w_2w_4}$, then the oriented angles $\angle\{[z_3z_2], [z_3z_1]\}$ and $\angle\{[z_4z_2], [z_4z_1]\}$ will be directed in different senses, and so their difference reduces to the *sum* of the angles z_3 and z_4 of the quadrangle; see Figure 58. Hence we obtain the following.

In order that a convex quadrangle $\overline{z_1z_3z_2z_4}$ *may be taken by a circular transformation into a convex quadrangle* $\overline{w_1w_3w_2w_4}$, *it is necessary and sufficient that the sum of the opposite angles* z_3 *and* z_4 *of the first quadrangle should be equal to the sum of the opposite*

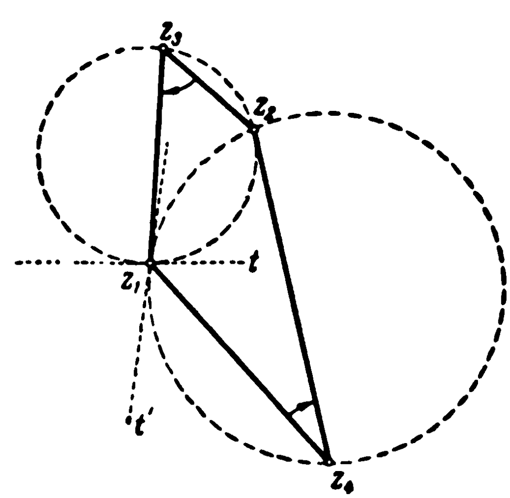


FIG. 58

angles w_3 and w_4 of the second quadrangle, and that the ratio $(z_1\,,\,z_3)(z_2\,,\,z_4)/(z_2\,,\,z_3)(z_1\,,\,z_4)$ of the products of opposite sides of the first quadrangle should be equal to the ratio

$$\frac{(w_1\,,\,w_3)(w_2\,,\,w_4)}{(w_2\,,\,w_3)(w_1\,,\,w_4)}$$

of the products of opposite sides of the second quadrangle.

From this proposition it follows, in particular, that *every convex quadrangle $\overline{z_1z_2z_3z_4}$ can be taken by a circular transformation into a parallelogram $\overline{z_1^0z_2^0z_3^0z_4^0}$, whose angles are equal to half the sums of opposite angles of the quadrangle $\overline{z_1z_2z_3z_4}$, and the squares of whose sides are the products of opposite sides of the quadrangle $\overline{z_1z_2z_3z_4}$.* In the case in which the quadrangle $\overline{z_1z_2z_3z_4}$ can be inscribed in a circle (in which the sums of its opposite angles are equal), this quadrangle can be taken by a circular transformation into a rectangle; if the products of its opposite sides are equal, it can be taken into a rhombus; finally, if it can be inscribed in a circle and the products of its opposite sides are equal (that is, it is **harmonic**; cf. p. 74), it can be taken by a circular transformation into a square [¶G]. Hence, it is not difficult to derive a whole series of various properties of quadrangles.

We note first of all that if the four points $z_1\,,\,z_2\,,\,z_3\,,$ and z_4 can be taken by a circular transformation into the four points $z_1^0\,,\,z_2^0\,,\,z_3^0\,,$ and $z_4^0\,,$ then

$$\begin{aligned}\frac{(z_1\,,\,z_2)(z_3\,,\,z_4)}{(z_1\,,\,z_3)(z_2\,,\,z_4)} &= \widetilde{W}(z_1\,,\,z_4\,,\,z_2\,,\,z_3) = \widetilde{W}(z_1^0\,,\,z_4^0\,,\,z_2^0\,,\,z_3^0)\\ &= \frac{(z_1^0\,,\,z_2^0)(z_3^0\,,\,z_4^0)}{(z_1^0\,,\,z_3^0)(z_2^0\,,\,z_4^0)}\end{aligned}$$

Therefore, for example, *if the quadrangle $\overline{z_1z_2z_3z_4}$ can be inscribed in a circle and the products of its opposite sides are equal, then the products of its diagonals is equal to twice the product of a pair of opposite sides* (since this is obviously true in the case of a square $\overline{z_1^0z_2^0z_3^0z_4^0}$); see Figure 59a.

The last proposition may be generalized to an arbitrary quadrangle $\overline{z_1z_2z_3z_4}$ which can be *inscribed in a circle*. For such a

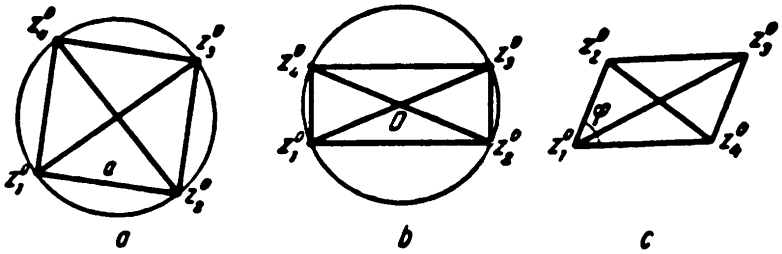


FIG. 59

quadrangle the sum of the products of opposite sides, $(z_1, z_2)(z_3, z_4) + (z_2, z_3)(z_1, z_4)$, *is equal to the product of the diagonals*, $(z_1, z_3)(z_2, z_4)$; this is **Ptolemy's theorem.** In fact, if the quadrangle $\overline{z_1z_2z_3z_4}$ can be taken by a circular transformation into a *rectangle* $\overline{z_1^0z_2^0z_3^0z_4^0}$ (Figure 59b), then

$$\frac{(z_1, z_2)(z_3, z_4)}{(z_1, z_3)(z_2, z_4)} + \frac{(z_2, z_3)(z_1, z_4)}{(z_1, z_3)(z_2, z_4)}$$
$$\doteq \tilde{W}(z_1, z_4, z_2, z_3) + \tilde{W}(z_1, z_2, z_4, z_3)$$
$$= \tilde{W}(z_1^0, z_4^0, z_2^0, z_3^0) + \tilde{W}(z_1^0, z_2^0, z_4^0, z_3^0)$$
$$= \frac{(z_1^0, z_2^0)(z_3^0, z_4^0)}{(z_1^0, z_3^0)(z_2^0, z_4^0)} + \frac{(z_2^0, z_3^0)(z_1^0, z_4^0)}{(z_1^0, z_3^0)(z_2^0, z_4^0)}$$
$$= \frac{(z_1^0, z_2^0)^2 + (z_2^0, z_3^0)^2}{(z_1^0, z_3^0)^2} = 1$$

Thus, Ptolemy's theorem can be regarded as a generalization of the theorem of Pythagoras, showing the connection between the lengths of the sides and the lengths of the diagonals of a rectangle.

We shall now try to discover an analogous relation connecting the lengths of the sides and diagonals of a completely arbitrary quadrangle $\overline{z_1z_2z_3z_4}$. Let this quadrangle be taken by a circular transformation into a parallelogram $\overline{z_1^0z_2^0z_3^0z_4^0}$ with acute angle $\angle z_1^0 = \varphi$; see Figure 59c. Then, obviously,

$$(z_2^0, z_4^0)^2 = (z_1^0, z_2^0)^2 + (z_1^0, z_4^0)^2 - 2(z_1^0, z_2^0)(z_1^0, z_4^0)\cos\varphi$$
$$(z_1^0, z_3^0)^2 = (z_1^0, z_2^0)^2 + (z_1^0, z_4^0)^2 + 2(z_1^0, z_2^0)(z_1^0, z_4^0)\cos\varphi$$

Consequently,

$$\begin{aligned}(z_1^0, z_3^0)^2(z_2^0, z_4^0)^2 \\ &= [(z_1^0, z_2^0)^2 + (z_1^0, z_4^0)^2]^2 - 4(z_1^0, z_2^0)^2(z_1^0, z_4^0)^2 \cos^2 \varphi \\ &= (z_1^0, z_2^0)^4 + (z_1^0, z_4^0)^4 - 2(z_1^0, z_2^0)^2(z_1^0, z_4^0)^2(2\cos^2 \varphi - 1) \\ &= (z_1^0, z_2^0)^4 + (z_1^0, z_4^0)^4 - 2(z_1^0, z_2^0)^2(z_1^0, z_4^0)^2 \cos 2\varphi\end{aligned}$$

But 2φ, the sum of opposite angles z_1^0 and z_3^0 of the parallelogram, is, as we know, equal to the sum of opposite angles of the quadrangle $\overline{z_1z_2z_3z_4}$. Hence we have

$$\begin{aligned}&\frac{(z_1, z_2)^2(z_3, z_4)^2 + (z_1, z_4)^2(z_2, z_3)^2}{(z_1, z_3)^2(z_2, z_4)^2} \\ &\qquad - \frac{2[(z_1, z_2)(z_3, z_4)][(z_1, z_4)(z_2, z_3)] \cos 2\varphi}{(z_1, z_3)^2(z_2, z_4)^2} \\ &= [\tilde{W}(z_1, z_4, z_2, z_3)]^2 + [\tilde{W}(z_1, z_2, z_4, z_3)]^2 \\ &\qquad - 2\tilde{W}(z_1, z_4, z_2, z_3)\tilde{W}(z_1, z_2, z_4, z_3) \cos 2\varphi \\ &= [\tilde{W}(z_1^0, z_4^0, z_2^0, z_3^0)]^2 + [W(z_1^0, z_2^0, z_4^0, z_3^0)]^2 \\ &\qquad - 2\tilde{W}(z_1^0, z_4^0, z_2^0, z_3^0)\tilde{W}(z_1^0, z_2^0, z_4^0, z_3^0) \cos 2\varphi \\ &= \frac{(z_1^0, z_2^0)^4 + (z_1^0, z_4^0)^4 - 2(z_1^0, z_2^0)^2(z_1^0, z_4^0)^2 \cos 2\varphi}{(z_1^0, z_3^0)^2(z_2^0, z_4^0)^2} = 1\end{aligned}$$

That is, *the sum of the products of the squares of opposite sides of an arbitrary quadrangle, minus twice the product of all the sides and the cosine of the sum of opposite angles, is equal to the product of the squares of the diagonals of the quadrangle.* In particular, if $2\varphi = \pi/2$ we obtain the result: *if the sum of opposite angles of a quadrangle is equal to $\pi/2$, then the sum of the products of the squares of opposite sides is equal to the product of the squares of the diagonals of the quadrangle* (compare with Ptolemy's theorem: *if the sum of opposite angles of a quadrangle is equal to π, then the sum of the products of opposite sides is equal to the product of the diagonals of the quadrangle*).

These results can also be arrived at in another way. By a circular transformation we take the quadrangle $\overline{z_1z_2z_3z_4}$ into a

"quadrangle" $\overline{Z_1Z_2Z_3\infty}$, where ∞ is the point at infinity of the plane. Then we have (see p. 34):

$$W(z_1, z_2, z_3, z_4) = W(Z_1, Z_2, Z_3, \infty) = V(Z_1, Z_2, Z_3)$$

The ratio $V(Z_1, Z_2, Z_3) = (Z_1 - Z_3)/(Z_2 - Z_3)$ gives the ratio of the sides $(Z_1, Z_3)/(Z_2, Z_3) = |Z_1 - Z_3|/|Z_2 - Z_3| = |V|$ of the triangle $\overline{Z_1Z_2Z_3}$ and its angle $\angle\{[Z_1Z_2], [Z_2Z_3]\} = \arg (Z_1 - Z_3)/(Z_2 - Z_3) = \arg V$; thus, the triangle $\overline{Z_1Z_2Z_3}$ determines the quadrangle $\overline{z_1z_2z_3z_4}$ "to within a similarity" (this expression means that any two triangles $\overline{Z_1Z_2Z_3}$ and $\overline{Z_1'Z_2'Z_3'}$, such that $\overline{z_1z_2z_3z_4}$ can be taken by circular transformations into $\overline{Z_1Z_2Z_3\infty}$ and into $\overline{Z_1'Z_2'Z_3'\infty}$, are similar). The triangle $\overline{Z_1Z_2Z_3}$ is called an **associated triangle** of the quadrangle $\overline{z_1z_2z_3z_4}$ (see p. 79).

Obviously,

$$\frac{(Z_1, Z_3)}{(Z_2, Z_3)} = |W(z_1, z_2, z_3, z_4)| = \tilde{W}(z_1, z_2, z_3, z_4)$$
$$= \frac{(z_1, z_3)(z_2, z_4)}{(z_1, z_4)(z_2, z_3)}$$

On the other hand,

$$\frac{(Z_1, Z_2)}{(Z_3, Z_2)} = \frac{|Z_1 - Z_2|}{|Z_3 - Z_2|} = |V(Z_1, Z_3, Z_2)|$$
$$= |W(Z_1, Z_3, Z_2, \infty) = |W(z_1, z_3, z_2, z_4)|$$
$$= \tilde{W}(z_1, z_3, z_2, z_4) = \frac{(z_1, z_2)(z_3, z_4)}{(z_1, z_4)(z_2, z_3)}$$

Thus, *the ratios of the lengths of the sides of an associated triangle* $\overline{Z_1Z_2Z_3}$ *are equal to the ratios of the products of opposite sides and diagonals of the quadrangle* $\overline{z_1z_2z_3z_4}$:

$$(Z_1, Z_2) : (Z_1, Z_3) : (Z_2, Z_3)$$
$$= [(z_1, z_2)(z_3, z_4)] : [(z_1, z_3)(z_2, z_4)] : [(z_1, z_4)(z_2, z_3)]$$

Hence it follows that if the quadrangle $\overline{z_1z_2z_3z_4}$ is taken by a circular transformation into a quadrangle $\overline{z_1'z_2'z_3'z_4'}$, where one of the points z_1', z_2', z_3', and z_4' is the point at infinity ∞, then the triangle formed by the other three points is associated with $\overline{z_1z_2z_3z_4}$ (that is, it is similar to the triangle $\overline{Z_1Z_2Z_3}$ considered above). In fact, if, for example, $z_3' = \infty$, then $\overline{z_1z_2z_4z_3}$ is taken into $\overline{z_1'z_2'z_4'\infty}$, whence it follows that

$$\begin{aligned}(z_1', z_2') &: (z_1', z_4') : (z_2', z_4') \\ &= [(z_1, z_2)(z_3, z_4)] : [(z_1, z_4)(z_2, z_3)] : [(z_1, z_3)(z_2, z_4)]\end{aligned}$$

and so the triangles $\overline{Z_1Z_2Z_3}$ and $\overline{z_1'z_2'z_4'}$ are similar (they have the same ratios of sides). It follows that an associated triangle of the quadrangle $\overline{z_1z_2z_3z_4}$ can be obtained, for example, by transforming any three vertices of $\overline{z_1z_2z_3z_4}$ by means of an inversion with center at the fourth vertex; see Figure 60.

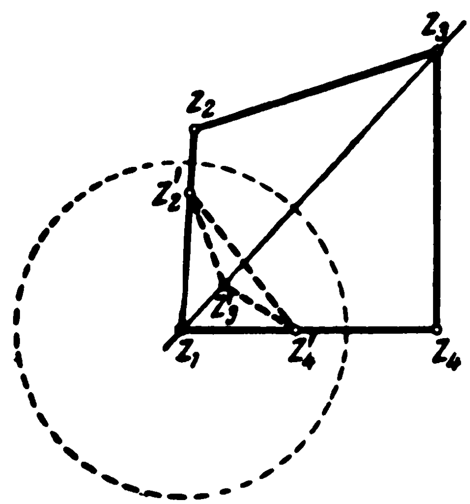


FIG. 60

We now note that if the quadrangle $\overline{z_1z_2z_3z_4}$ is convex, then (see pp. 34 and 146):

$$\begin{aligned}\angle\{[Z_2Z_3], [Z_2Z_1]\} &= \arg V(Z_1, Z_3, Z_2) \\ &= \arg W(Z_1, Z_3, Z_2, \infty) \\ &= \arg W(z_1, z_3, z_2, z_4) \\ &= \angle z_1 + \angle z_3 = 2\varphi\end{aligned}$$

Hence it follows that the relation derived above,

$$(z_1\,,\,z_3)^2(z_2\,,\,z_4)^2=(z_1\,,\,z_2)^2(z_3\,,\,z_4)^2+(z_1\,,\,z_4)^2(z_2\,,\,z_3)^2$$
$$-\,2(z_1\,,\,z_2)(z_2\,,\,z_3)(z_3\,,\,z_4)(z_4\,,\,z_1)\cos 2\varphi$$

simply consists of the **cosine theorem** for the associated triangle. In the same way, the theorem relating to the case $2\varphi=90°$ and Ptolemy's theorem can easily be derived by considering an associated triangle of the quadrangle $\overline{z_1z_2z_3z_4}$ (right-angled or degenerate respectively).

One further set of results connected with quadrangles can be obtained because of the possibility of taking any quadrangle into a parallelogram (or rhombus or square). The "circles" drawn through opposite vertices of a rectangle $\overline{z_1^0z_2^0z_3^0z_4^0}$ and **orthogonal** to the circumscribing circle (circles are orthogonal if the angle between them is $\pi/2$) are the diagonals $[z_1^0z_3^0]$ and $[z_2^0z_4^0]$ of the rectangle; the points of intersection of these circles are the center O of the rectangle and the point ∞ (through which all lines of the plane pass). Further, the circles $[z_1^0z_2^0\infty]$, $[z_2^0z_3^0\infty]$, $[z_3^0z_4^0\infty]$, and $[z_4^0z_1^0\infty]$ are the sides of the rectangle; it is easy to imagine what the circles $[z_1^0z_2^0O]$, $[z_2^0z_3^0O]$, $[z_3^0z_4^0O]$, and $[z_4^0z_1^0O]$ look like (see Figure 59b). Hence we obtain the following result: *If s_1 and s_2 are the circles which pass through pairs of opposite vertices of a quadrangle $\overline{z_1z_2z_3z_4}$ inscribed in a circle S and are orthogonal to S, and if v and w are the points of intersection of these circles, then the circles $[z_1z_2v]$ and $[z_3z_4v]$, $[z_1z_4v]$ and $[z_2z_3v]$ touch each other, and the first two circles are orthogonal to the last two; similarly, the circles $[z_1z_2w]$ and $[z_3z_4w]$, $[z_1z_4w]$ and $[z_2z_3w]$ touch each other, and the first two of these circles are orthogonal to the last two*; see Figure 61. Further, the fact that the point O is the midpoint of the diagonals $\overline{z_1^0z_3^0}$ and $\overline{z_2^0z_4^0}$ of the rectangle $\overline{z_1^0z_2^0z_3^0z_4^0}$ leads to some curious propositions relating to an arbitrary quadrangle $\overline{z_1z_2z_3z_4}$ which can be inscribed in a circle, for it follows from this that

$$\widetilde{W}(z_1\,,\,z_3\,,\,v,\,w)=\widetilde{W}(z_2\,,\,z_4\,,\,v,\,w)=1$$

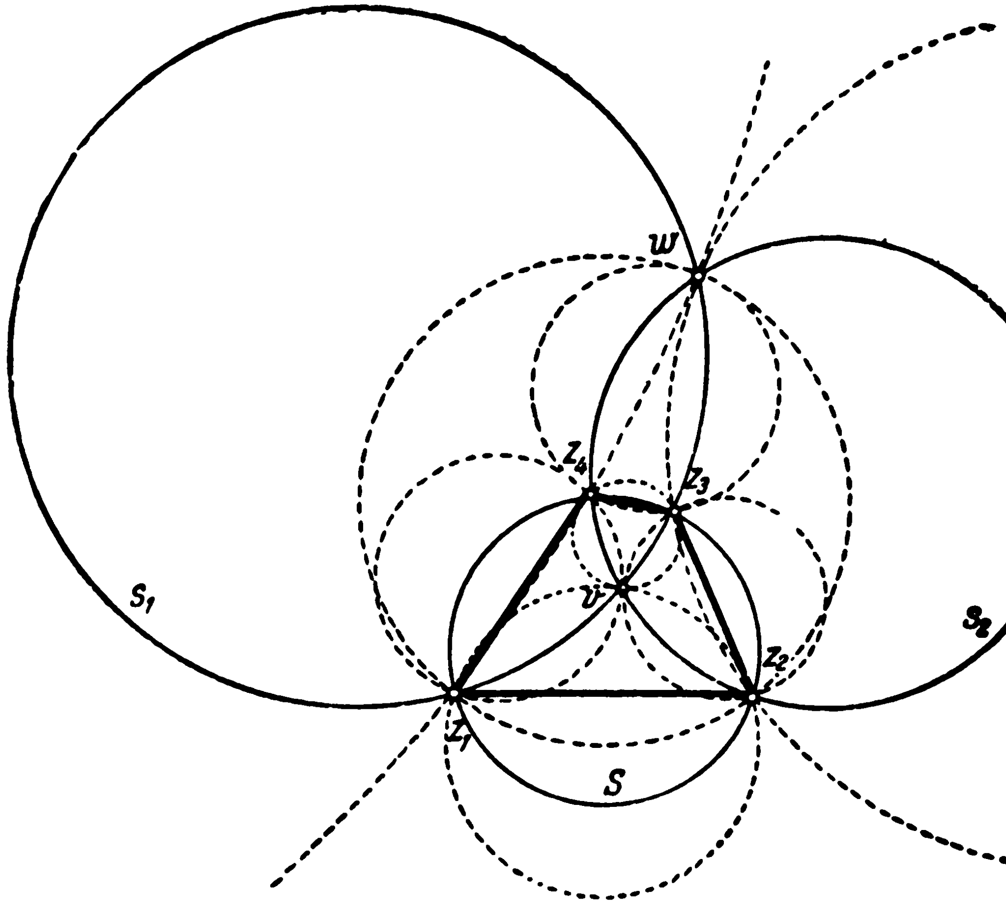


FIG. 61

or, put another way,

$$\frac{(z_1\,,\,v)}{(z_3\,,\,v)}=\frac{(z_1\,,\,w)}{(z_3\,,\,w)}\qquad\text{and}\qquad\frac{(z_2\,,\,v)}{(z_4\,,\,v)}=\frac{(z_2\,,\,w)}{(z_4\,,\,w)}$$

It follows from Figure 59a that *if the products of opposite sides of a quadrangle* $\overline{z_1z_2z_3z_4}$ (inscribed in a circle S) *are equal, then the circles* s_1 *and* s_2 *are orthogonal*; moreover, *the circles* $[z_1z_2v]$, $[z_2z_3v]$, $[z_3z_4v]$, *and* $[z_4z_1v]$ *are orthogonal to the circles* $[z_1z_2w]$, $[z_2z_3w]$, $[z_3z_4w]$, *and* $[z_4z_1w]$, *respectively*. Finally, many of these results can be carried over to a completely arbitrary quadrangle $\overline{z_1z_2z_3z_4}$ (not necessarily inscribable in a circle); here we must understand by s_1 and s_2, respectively, the circle which passes through the opposite vertices z_1 and z_3 of the quadrangle and makes equal angles with the circles $[z_1z_2z_3]$ and $[z_1z_3z_4]$ and the circle which passes through z_2 and z_4 and makes equal angles with $[z_2z_3z_4]$ and $[z_2z_4z_1]$; if the original quadrangle is a paral-

lelogram, then s_1 and s_2 are the diagonals of the parallelogram (see Figure 59c). We leave the reader to consider this case for himself.

In conclusion, we note that the examples just considered give a good illustration of the method of proving theorems of circular geometry, of which we spoke at the end of the preceding section. The significance of our discussion lies in the fact that we formulate a certain *circular property* of a quadrangle $\overline{z_1z_2z_3z_4}$, which is a property preserved by circular transformations (for example, one connected with the cross-ratio of the vertices of the quadrangle), and we then transform the quadrangle into such a form that the property is discovered more easily for this form than for the original quadrangle (for example, by taking $\overline{z_1z_2z_3z_4}$ into a parallelogram $\overline{z_1^0z_2^0z_3^0z_4^0}$ or a quadrangle $\overline{Z_1Z_2Z_3\infty}$). This method enables us to prove many different theorems relating to polygons and circles. Here we shall limit ourselves to just one very simple example.

We consider four nonintersecting circles S_1, S_2, S_3, and S_4, such that S_1 *and* S_3 *touch* S_2 *and* S_4; see Figure 62a. *It is required to show that the four points of contact* (denoted by z_1, z_2, z_3, and z_4) *lie on one circle* (*or line*) Σ.[73] By a circular transformation we take the point z_4 into the point at infinity ∞. In this case the

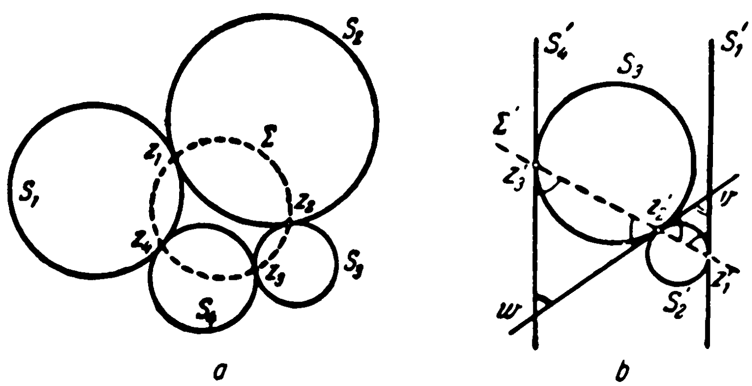


FIG. 62

[73] If the circles S_1 and S_3, or S_2 and S_4, intersect, the conditions of the theorem cannot hold. (It is possible to dispense with the requirement that the circles do not intersect if it is supposed that they are *oriented*; cf. footnote 81.)

circles S_1 and S_4, which pass through z_4, go into parallel lines S'_1 and S'_4 (S'_1 and S'_4 cannot intersect, because S_1 and S_4 have the *unique* common point z_4); see Figure 62b. From Figure 62b it follows that the points z'_1, z'_2, and z'_3, into which the points z_1, z_2, and z_3 go, lie on one line Σ'; in fact, $\angle(z'_1 v z'_2) = \angle(z'_2 w z'_3)$, and so $\angle(z'_1 z'_2 v) = \angle(z'_3 z'_2 w)$, and hence the points z'_1, z'_2, and z'_3 lie on one line. It follows that the number $V(z'_1, z'_2, z'_3) = W(z'_1, z'_2, z'_3, \infty)$ is real, and so the cross-ratio $W(z_1, z_2, z_3, z_4)$ is real; but this proves the theorem.

We now investigate the connection between the transformation of inversion and the idea of the power of a point with respect to a circle, introduced in Section 8. In Section 8 it was shown that the ratio

$$C/A = k \tag{18}$$

of the coefficients C and A in the equation of a circle S,

$$Az\bar{z} + Bz - \bar{B}\bar{z} + C = 0, \qquad A, C \text{ purely imaginary} \tag{19}$$

is equal to the product

$$\{O, z_1\} \cdot \{O, z_2\} \tag{20}$$

of the (oriented) lengths of the segments $\overrightarrow{Oz_1}$ and $\overrightarrow{Oz_2}$, where z_1 and z_2 are the points of intersection of S and an arbitrary line passing through O; we called this product the **power of the circle** S (or the *power of the point* O *with respect to the circle* S). On the other hand, the inversion (Equation 16) of power k

$$z' = \frac{k}{\bar{z}}, \qquad k \text{ real}$$

takes each point z into the point z' of the line $[Oz]$ such that

$$\{O, z\} \cdot \{O, z'\} = k$$

See the geometrical description of the unit inversion, p. 139. It follows immediately that, *if the power of the point* O *with respect to the circle* S *is equal to* k, *then the inversion of power* k

takes S into itself (it takes each point z_1 of the circle S into the second point of intersection z_2 of the line $[Oz_1]$ and S).

The fact that a circle (Equation 19), such that $C/A = k$, goes into itself under an inversion of power k is not difficult to prove without resorting to the idea of the power of a point with respect to a circle. On the contrary, from this proof it may be deduced that the product, $\{O, z_1\} \cdot \{O, z_2\}$ is constant, where z_1 and z_2 are the points of intersection of our circle S and a line passing through O (the **theorem of the power of a point with respect to a circle**). In fact, the inversion, Equation 16, which may also be written in the forms

$$z = \frac{k}{\bar{z}'}, \qquad \bar{z} = \frac{k}{z'}$$

takes the circle, Equation 19, into the locus of points z', for which

$$A(k/\bar{z}')(k/z') + B(k/\bar{z}') - \bar{B}(k/z') + C = 0$$

or

$$Cz'\bar{z}' + Bkz' - \bar{B}k\bar{z}' + Ak^2 = 0$$

Since $C = Ak$ and $A = C/k$, the latter equation can be written

$$Akz'\bar{z}' + Bkz' - \bar{B}k\bar{z}' + Ck = 0$$

from which it follows that the inversion, Equation 16, takes the circle, Equation 19, into itself.

For each inversion, Equation 16, there exist *infinitely many* circles which go into themselves under this inversion. These are circles of power k; they are given by Equation 19 with C/A fixed, that is by the equation

$$Az\bar{z} + Bz - \bar{B}\bar{z} + Ak = 0 \tag{19a}$$

The lines $Bz - \bar{B}\bar{z} = 0$ passing through O, which also go into themselves under the inversion of Equation 16, are given by this last equation (where we just have to put $A = 0$). The set of circles (and lines) given by this last equation, which are taken

into themselves by the inversion (Equation 16) of power k [¶H], is called a **bundle of circles**; the number k is called the **power of the bundle**, and the point O is called its **center**. It may be shown that every bundle of circles consists of all circles (and lines) which intersect some fixed circle Σ at right angles, or of all circles (and lines) which pass through some fixed point, or of all circles (and lines) which intersect some fixed circle Σ at diametrically opposite points [¶I].[74]

§15. Axial Circular Transformations (Laguerre Transformations)

In this section we shall consider linear-fractional functions of a **dual variable**, Equations 1 and 1a,

$$z' = \frac{az + b}{cz + d}, \qquad z' = \frac{a\bar{z} + b}{c\bar{z} + d}$$

where we must now suppose that the determinant $\Delta = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$ is not a divisor of zero. To these functions correspond transformations of the set of oriented lines (axes) of the Euclidean plane, which we shall call **axial linear-fractional transformations**.[75] As before, it is sometimes convenient to call Equation 1 an **axial homography** and Equation 1a an **axial antihomography**. These linear-fractional functions of a dual variable are single-valued functions, defined on the set of all dual numbers, which are extended by introducing the numbers $c\omega$, where c is real, and ∞;[76] correspondingly, the axial linear-fractional transformations are in one-to-one correspondence with the transformations of the set of all axes (oriented lines) of the plane.

---

[74] See, for example, H. Schwerdtfeger, *Geometry of Complex Numbers*, §4 (Toronto University Press, Oliver and Boyd), 1962.

[75] In contrast to these transformations, the linear-fractional transformations of the set of points of the plane considered in Sections 13 and 14 can be called *linear-fractional point transformations*.

[76] See above, pp. 15–16.

Particular cases of axial linear-fractional transformations are the *symmetry about the point O*, the *symmetry about the line o*, and the *reorientation*,

$$z' = \bar{z}, \qquad z' = -z, \qquad \text{and} \qquad z' = -\frac{1}{\bar{z}} \tag{21a,b,c}$$

and also (see Equations 36a-d of Chapter II) the arbitrary **motions**,

$$z' = \frac{pz+q}{-\bar{q}z+\bar{p}}, \qquad z' = \frac{-pz+q}{\bar{q}z+\bar{p}}, \qquad z' = \frac{p\bar{z}+q}{-\bar{q}\bar{z}+\bar{p}},$$

$$z' = \frac{-p\bar{z}+q}{\bar{q}\bar{z}+\bar{p}}, \qquad \Delta = p\bar{p}+q\bar{q} \neq 0 \tag{22a,b,c,d}$$

Just as in Section 13, it may be shown that *the product of two axial linear-fractional transformations and the inverse of a linear-fractional transformation are also transformations of the same type*; finally, *the identical transformation can be regarded as axial linear-fractional.* It is very important to note that, *if z_1', z_2', z_3', and z_4' are the four* (oriented) *lines of the plane into which the axial linear-fractional transformation* (Equation 1 or 1a) *takes four given lines z_1, z_2, z_3, and z_4, then we have*

$$W(z_1', z_2', z_3', z_4') = W(z_1, z_2, z_3, z_4)$$

or

$$W(z_1', z_2', z_3', z_4') = \overline{W(z_1, z_2, z_3, z_4)}$$

which is the **property of the invariance of the cross-ratio** (Equations 9 and 9a); the proof of this property is exactly the same as the proof of the corresponding statement in Section 13. Hence it follows immediately that *axial linear-fractional transformations take four* (oriented) *lines, which touch one* (oriented) *circle or pass through one point, into four lines, which also touch one circle or pass through one point*; in other words, *axial linear-fractional transformations take each* (oriented) *circle or point into a circle or point.*[77] This enables us to call axial linear-fractional

[77] Cf. footnote 64.

transformations of the plane **axial circular transformations**, where the transformations such as Equation 1 are called **direct** axial circular transformations and those such as Equation 1a are called **opposite** axial circular transformations.[78] Since axial circular transformations were first considered by the outstanding French mathematician E. Laguerre (1834–1886), they are often called **Laguerre transformations**.

As in Section 13, it may be shown that *there exists a unique axial homography* (Equation 1) *and a unique axial antihomography* (Equation 1a), *each taking three given* (oriented) *lines* z_1, z_2, *and* z_3, no two of which are parallel, *into three other given* (oriented) *lines* w_1, w_2, *and* w_3, no two of which are parallel.[79] These transformations are expressed by Equations 10 and 10a:

$$\frac{z' - w_2}{w_1 - w_2} : \frac{z' - w_3}{w_1 - w_3} = \frac{z - z_2}{z_1 - z_2} : \frac{z - z_3}{z_1 - z_3},$$

$$\frac{z' - w_2}{w_1 - w_2} : \frac{z' - w_3}{w_1 - w_3} = \frac{\bar{z} - \bar{z}_2}{\bar{z}_1 - \bar{z}_2} : \frac{\bar{z} - \bar{z}_3}{\bar{z}_1 - \bar{z}_3}$$

---

[78] It may be shown that *all* transformations of the set of oriented lines of the plane which take (oriented) circles (including points) into circles are exhausted by the linear-fractional transformations (Equations 1 and 1a), possibly combined with a similarity (see I. M. Yaglom, *Geometric Transformations*, Part II, Chap II, Section 5 (Randoni House, New York), 1966).

[79] It is not difficult to see that *every axial circular transformation takes parallel lines into parallel lines*. This follows, for example, from the fact that, by virtue of Equations 1 and 1a, $|z'| = (|a||z| + |b|)/(|c||z| + |d|)$, and so, if $|z_1| = |z_2|$, then $|z_1'| = |z_2'|$. Therefore, if the lines z_1 and z_2 are parallel, then w_1 and w_2 must be parallel (if $z_1 - z_2$ is a divisor of zero, but $w_1 - w_2$ is not a divisor of zero, then the determinant Δ of the linear-fractional transformation given by Equation 10 is a divisor of zero). In general, *in order that there may exist a* (direct) *axial circular transformation which takes three lines* z_1, z_2, *and* z_3 *into three lines* w_1, w_2, *and* w_3, *it is necessary that no two of the lines* z *and no two of the lines* w *should be parallel, or that two lines* z *and the two lines* w *corresponding to them should be parallel and the third line* z *and the third line* w *should not be parallel to the first two, or that* z_1, z_2, *and* z_3 *should be parallel to each other, and* w_1, w_2, *and* w_3 *parallel to each other, and* $\{z_1, z_2\} : \{z_1, z_3\}$ *equal to* $\{w_1, w_2\} : \{w_1, w_3\}$.

On the other hand, four (oriented) lines z_1, z_2, z_3, and z_4 cannot *always* be taken, by an axial circular transformation, into four other (oriented) lines w_1, w_2, w_3, and w_4. In order that this may be possible, *it is necessary and sufficient that one of the following* (Equation 9 or 9a) *hold*:

$$W(w_1, w_2, w_3, w_4) = W(z_1, z_2, z_3, z_4)$$

or

$$W(w_1, w_2, w_3, w_4) = \overline{W(z_1, z_2, z_3, z_4)}$$

From what has been proved it follows that *any* (oriented) *circle or point can* (in many ways) *be taken by an axial circular transformation into any other circle or point*; for this it is merely necessary that three (oriented) tangents of the first circle should go into any three oriented tangents of the second circle. In particular, *any* (oriented) *circle can be taken, by an axial circular transformation, into a point*; by virtue of this, in problems connected with these transformations it is usual not to distinguish between points and circles, regarding a point as a special case of a circle, a "circle of zero radius".

We now explain the geometrical significance of the argument arg W and the modulus $|W|$ of the cross-ratio $W(z_1, z_2, z_3, z_4) = (z_1 - z_3)/(z_2 - z_3) : (z_1 - z_4)/(z_2 - z_4)$ of the four (oriented) lines z_1, z_2, z_3, and z_4. In Section 9 we proved in effect that

$$\begin{aligned} \arg W(z_1, z_2, z_3, z_4) &= \arg \frac{z_1 - z_3}{z_2 - z_3} - \arg \frac{z_1 - z_4}{z_2 - z_4} \\ &= \tfrac{1}{2}(\{[z_1z_3], [z_2z_3]\} + \{[z_2z_4], [z_1z_4]\} \\ &\quad - \{[z_4z_1], [z_3z_1]\} - \{[z_3z_2], [z_4z_2]\}) \end{aligned}$$

We now consider two (oriented) circles S_1 and S_2, determined by the lines z_1, z_2, z_3, and z_1, z_2, z_4, or, as we shall often denote them subsequently, the circles $[z_1z_2z_3]$ and $[z_1z_2z_4]$; see Figure 63. We denote the points of contact of the circles S_1 and S_2 with the lines z_1, z_2, z_3 and z_1, z_2, z_4 by P_1, P_2, P_3 and Q_1, Q_2, Q_4, respectively; in addition, let us agree for simplicity to write A,

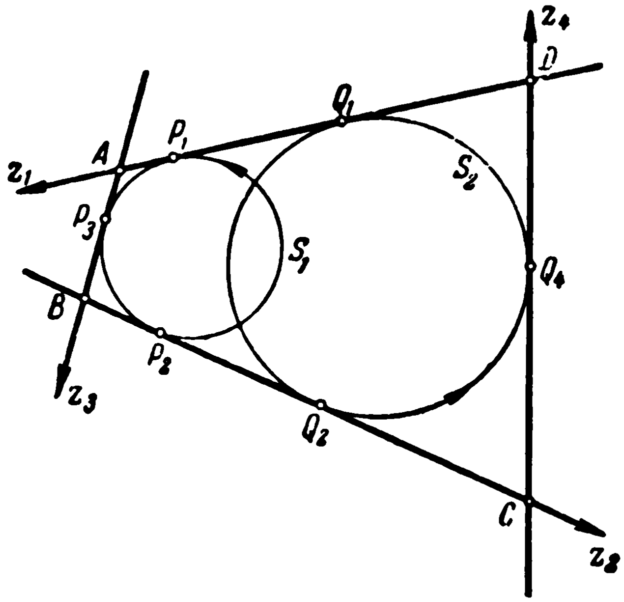


FIG. 63

B, C, and D instead of $[z_1z_3]$, $[z_2z_3]$, $[z_2z_4]$, and $[z_1z_4]$. Obviously we have (cf. pp. 91–92):

$$\begin{aligned}
&\tfrac{1}{2}(\{A, B\} + \{C, D\} - \{D, A\} - \{B, C\}) \\
&\quad = \tfrac{1}{2}(\{A, P_3\} + \{P_3, B\} + \{C, Q_4\} + \{Q_4, D\} \\
&\qquad - \{D, Q_1\} - \{Q_1, P_1\} - \{P_1, A\} \\
&\qquad - \{B, P_2\} - \{P_2, Q_2\} - \{Q_2, C\})
\end{aligned}$$

Further, by virtue of the properties of tangents to circles,

$$\{A, P_3\} = \{P_1, A\}, \qquad \{P_3, B\} = \{B, P_2\}, \qquad \{C, Q_4\} = \{Q_2, C\},$$
$$\{Q_4, D\} = \{D, Q_1\}$$
$$\{P_2, Q_2\} = \{Q_1, P_1\} = -\{P_1, Q_1\}$$

Hence we obtain

$$\arg W(z_1, z_2, z_3, z_4) = \{P_1, Q_1\} \tag{23}$$

The length of the segment of the common tangent z to two (oriented) circles S_1 and S_2, included between the points of contact, is called the **tangential distance** of these circles and is denoted by (S_1, S_2); if the length of the segment of the common

tangent to the circles S_1 and S_2 is regarded as oriented, we speak of the **oriented tangential distance** $\{S_1 z S_2\}$ of these circles. Thus we see that *the argument* $\arg W$ *of the cross-ratio* $W(z_1, z_2, z_3, z_4)$ *of the four* (oriented) *lines* z_1, z_2, z_3, *and* z_4 *is equal to the* (oriented) *tangential distance* $\{[z_1z_2z_3]z_1[z_1z_2z_4]\}$ *of the circles* $[z_1z_2z_3]$ *and* $[z_1z_2z_4]$.

From the property of the invariance of the cross-ratio of four (oriented) lines it follows that *a direct axial circular transformation does not alter the oriented tangential distance of circles*, and *an opposite axial circular transformation alters the sign, but not the absolute value, of this distance*. This important property of axial circular transformations is usually formulated thus: *axial circular transformations preserve tangential distances of circles*; see Figure 64.[80] In particular, *under axial circular transformations touching* (oriented) *circles* (circles whose tangential distance is equal to zero) *go into touching circles* [¶J].[81]

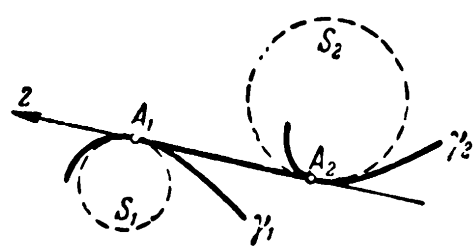


FIG. 64

[80] Let γ_1 and γ_2 be two arbitrary curves and z a common tangent of them (Figure 64); then the distance $(A_1, A_2) = (\gamma_1, \gamma_2)$ between the points of contact A_1 and A_2 of γ_1 and γ_2 with z is called the **tangential distance** of γ_1 and γ_2. But obviously $(\gamma_1, \gamma_2) = (S_1, S_2)$, where S_1 and S_2 are circles which touch the curves γ_1 and γ_2 at the points A_1 and A_2; furthermore, the axial circular transformation which takes the curves γ_1 and γ_2 into other curves γ_1' and γ_2' takes S_1 and S_2 into circles S_1' and S_2' which touch γ_1' and γ_2'. It follows, therefore, that *axial circular transformations preserve tangential distances of arbitrary curves*. The transformations of the set of curves of the plane which have this property are called **equidistantial transformations**; thus, *axial circular transformations are equidistantial transformations*.

[81] We note that oriented circles are called **touching** if they have a unique common oriented tangent, that is, if they touch in the usual

Let us now turn to the modulus $|W|$ of the cross-ratio $W(z_1, z_2, z_3, z_4)$ of four (oriented) lines z_1, z_2, z_3, and z_4. Using the fundamental formula, Equation 29 of Chapter II, and the fact that the modulus of the quotient or difference of two dual numbers is equal to the quotient or difference, respectively, of the moduli of these numbers, we obtain

$$| W(z_1, z_2, z_3, z_4)| = \frac{\tan \angle\{z_1, o\}/2 - \tan \angle\{z_3, o\}/2}{\tan \angle\{z_2, o\}/2 - \tan \angle\{z_3, o\}/2} : \frac{\tan \angle\{z_1, o\}/2 - \tan \angle\{z_4, o\}/2}{\tan \angle\{z_2, o\}/2 - \tan \angle\{z_4, o\}/2}$$

Taking into account the fact $\tan \alpha - \tan \beta = \sin (\alpha - \beta)/(\cos \alpha \cos \beta)$; we have

$$| W(z_1 \; z_2, z_3, z_4) | = \frac{\sin(\frac{1}{2}\angle\{z_3, z_1\})}{\sin(\frac{1}{2}\angle\{z_3, z_2\})} : \frac{\sin(\frac{1}{2}\{z_4, z_1\})}{\sin(\frac{1}{2}\{z_4, z_2\})}$$

We shall call the real number $\sin(\frac{1}{2}\{z_3, z_1\})/\sin(\frac{1}{2}\{z_3, z_2\})$: $\sin(\frac{1}{2}\{z_4, z_1\})/(\sin(\frac{1}{2}\{z_4, z_2\})$ the **cross-ratio of the angles between the four** (oriented) **lines** z_1, z_2, z_3, and z_4, and denote it by $\hat{W}(z_1, z_2, z_3, z_4)$; thus,

$$| W(z_1, z_2, z_3, z_4)| = \hat{W}(z_1, z_2, z_3, z_4) \tag{24}$$

From the invariance of the cross-ratio of four (oriented) lines under axial circular transformations we may conclude that *axial circular transformations preserve the cross-ratio of the angles between four (oriented) lines.* We may now say that *four* (oriented) *lines z_1, z_2, z_3, and z_4 can be taken by an axial circular transformation into four other lines w_1, w_2, w_3, and w_4, if and only if the tangential distance of the circles $[z_1z_2z_3]$ and $[z_1z_2z_4]$ is equal to the tangential distance of the circles $[w_1w_2w_3]$ and $[w_1w_2w_4]$ and the cross-ratio of the angles between z_1, z_2, z_3, and z_4 is equal to the cross-ratio of the angles between w_1, w_2, w_3, and w_4.*

---

sense and their directions at the point of contact coincide (circles touching in the usual sense, whose directions at the point of contact are opposite, can be taken by axial circular transformations into intersecting circles or circles with no common points).

We now touch on the question of the geometrical description of all axial circular transformations. Some of the simplest transformations of this kind, apart from the motions given by Equations 22, are the transformations

$$z' = k/z \qquad \text{and} \qquad z' = k/\bar{z}, \qquad k \text{ real} \tag{25a,b}$$

If k is not equal to ± 1,[82] these transformations can be rewritten as

$$\begin{aligned} \arg z' &= -\arg z, \qquad | z' | = k/| z | \\ \arg z' &= \arg z, \qquad | z' | = k/| z | \end{aligned} \tag{26a,b}$$

The simpler of these two transformations is Equation 25b or 26b, which is called an **axial inversion of power** k; the line o is called the **axis** of this inversion. Under an axial inversion of power k each (oriented) line z of the plane goes into a line z', *which meets the axis of inversion o in the same point M as z does, and is such that*

$$\tan \frac{\angle\{z, o\}}{2} \cdot \tan \frac{\angle\{z', o\}}{2} = k \tag{27}$$

See Figure 65; the line z' obviously goes into z. A line $t = (p/2)\varepsilon$ parallel to o and separated from o by an (oriented) distance $\{o, t\} = p$ is taken by the axial inversion (Equation 25b)

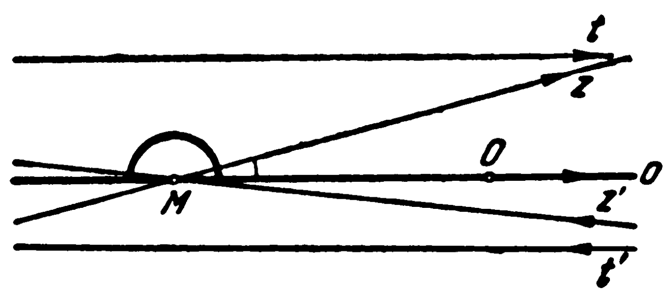


FIG. 65

[82] If $k = \pm 1$ the transformations (Equations 25) are obviously motions; they are special cases of transformations (Equations 22) where $p = 0$ and $q = 1$; thus, the transformation $z' = -1/\bar{z}$ represents the reorientation (see Equation 21c); the transformation $z' = 1/\bar{z}$ represents the symmetry about the axis o, together with a reorientation (the product of the transformations given by Equations 21b and 21c; in certain respects it is convenient to call this transformation $z' = 1/\bar{z}$ the "symmetry about the axis o" in the domain of directed lines).

into a line $t' = k/t = -(2k/p)\omega$, antiparallel to o and separated from o by a distance $\{o, t'\} = p/k$ (and t' is taken into t); in particular, the transformation takes the axis o into a line o_1, differing from o only in direction, and it takes o_1 into o. As for the transformation given by Equation 25a, it is a combination of the axial inversion of Equation 25b and the symmetry about the origin O of Equation 21a.

As well as the motions given by Equations 22,

$$z' = \frac{pz + q}{-\bar{q}z + \bar{p}}, \qquad z' = \frac{-pz + q}{\bar{q}z + \bar{p}},$$

$$z' = \frac{p\bar{z} + q}{-\bar{q}\bar{z} + \bar{p}}, \qquad \text{or} \qquad z' = \frac{-p\bar{z} + q}{\bar{q}\bar{z} + \bar{p}}$$

and the inversion given by Equation 25b,

$$z' = k/\bar{z}$$

there is one interesting axial circular transformation which merits special attention, and which may be compared with the translation in the direction perpendicular to the polar axis $z' = (z + q)/(qz + 1)$, where $q = (t/2)\varepsilon$ (see Equation 32a of Chapter II); this transformation is

$$z' = \frac{z + q}{-qz + 1}, \qquad q = \varepsilon\frac{t}{2}, \qquad |\, q \,| = 0 \tag{28}$$

The geometrical significance of this transformation is very simple: it takes each line $z = \tan\theta/2 \cdot (1 + \varepsilon s)$ into a line

$$\begin{aligned} z' &= \tan\frac{\theta'}{2}(1 + \varepsilon s') \\ &= \left(\tan\frac{\theta}{2}(1 + \varepsilon s) + \varepsilon\frac{t}{2}\right) : \left(-\tan\frac{\theta}{2}(1 + \varepsilon s)\varepsilon\frac{t}{2} + 1\right) \\ &= \tan\frac{\theta}{2}\left[1 + \varepsilon\left(s + \frac{t}{2}\cot\frac{\theta}{2}\right)\right] : \left(1 - \varepsilon\frac{t}{2}\tan\frac{\theta}{2}\right) \\ &= \tan\frac{\theta}{2}\left\{1 + \varepsilon\left[s + \frac{t}{2}\left(\tan\frac{\theta}{2} + \cot\frac{\theta}{2}\right)\right]\right\} \\ &= \tan\frac{\theta}{2}\left[1 + \varepsilon\left(s + \frac{t}{\sin\theta}\right)\right] \end{aligned}$$

parallel to z (since $|\, z' \,| = |\, z \,|$) and such that the distances

p and p' of the lines z and z' from the pole O of the system of polar coordinates are connected by the relation (see Equation 30 of Chapter II):

$$p' = s' \sin \theta' = \left(s + \frac{t}{\sin \theta}\right) \sin \theta = s \sin \theta + t = p + t \tag{29}$$

In other words, *the line z' is parallel to the line z, and the distance $\{z, z'\}$ from the line z to z' is equal to t*; see Figure 66. Equation 28

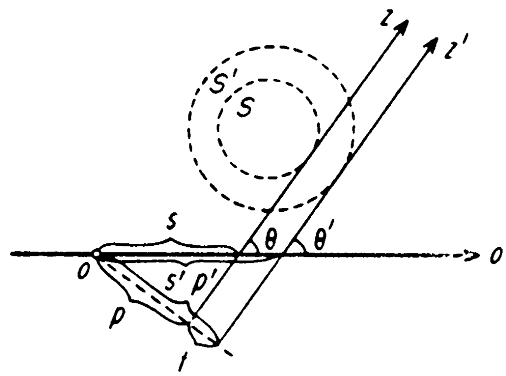


FIG. 66

is called an **axial dilatation** of (positive or negative) magnitude t.[83] It is obvious that the dilatation takes an (oriented) circle S of (positive or negative) radius r into a circle S' concentric with S and of radius $r + t$; in particular, it takes a point into a circle of radius t and a circle of radius $-t$ into a point.

We may show that *any axial circular transformation* (Equation 1 or 1a) *other than a motion* (Equations 22) *consists of the product of a motion and an axial inversion or the product of a motion and a dilatation* [¶K]. It would be possible to verify this by direct calculation, as we did in Section 13, by choosing a motion and an axial inversion (or a motion and a dilatation) whose product gives the preassigned axial circular transformation (Equation 1 or 1a). However, since this method leads to some complicated calculation, we shall proceed in a different way.

We note first of all that *an axial circular transformation*

---

[83] This axial dilatation should not be confused with the dilatation (or "point dilatation") defined earlier (beginning of Section 7) as a point transformation. The word "axial" will be understood in the rest of this section.—TRANSL.

(Equation 1 *or* 1a) *can be obtained by means of a motion followed by an axial inversion, if and only if this transformation takes one pair of lines* z *and* z_1 *differing only in direction into lines* z' *and* z_1' *also differing only in direction.* In fact, let an axial circular transformation consist of the product of a motion and an axial inversion with axis z_1' ; then this transformation takes two lines differing only in direction into the lines z_1' and z', where $z' = -1/\bar{z}_1'$ differs from z_1' in direction. Let us consider now an arbitrary axial homography (Equation 1) with the property that it takes the lines z and $z_1 = -1/\bar{z}$ into the lines z' and $z_1' = -1/\bar{z}'$. Let this transformation take a line z_2 not parallel to z or z_1 into a line z_2' (see Figure 67), where z_2' is not parallel to z' or z_1', since only

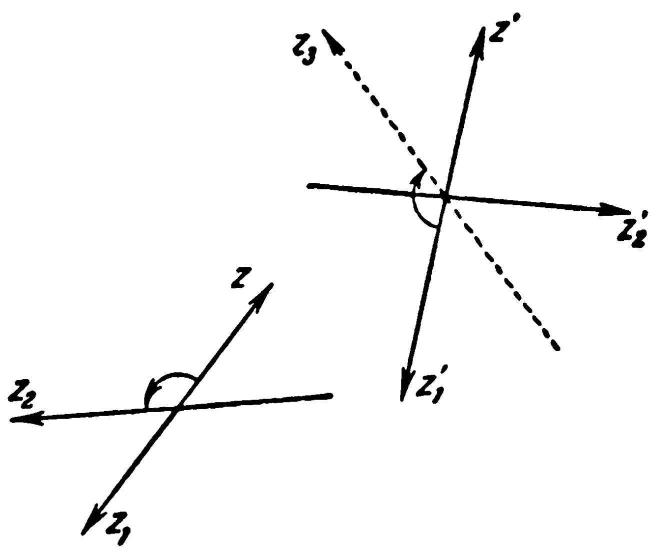


FIG. 67

parallel lines can be taken into parallel lines by an axial circular transformation.[84] We assert that our transformation consists of the product of an opposite motion (Equations 22c,d), which takes z into z_1' and z_1 into z' and the line z_2 into some line z_3 meeting z_1' in the same point as z_3' does, and the axial inversion with axis z_1' and power given by:[85]

$$k = \tan \frac{\angle\{z_1', z_2'\}}{2} \cdot \tan \frac{\angle\{z_1', z_3'\}}{2}$$

[84] See footnote 79.

[85] If the original transformation is itself a motion, then $k = \pm 1$.

In fact, this product of a motion and an axial inversion takes z into z', z_1 into z_1', and z_2 into z_2', and there exists *only one* axial homography which takes three given nonparallel lines into three known lines. In the same way it may be shown that, if an axial antihomography (Equation 1a) takes a pair of lines differing only in direction into lines differing only in direction, then it can be represented as the product of a *direct* motion (Equations 22a,b) and an inversion.

We show now that *each axial homography* (Equation 1) *can be expressed as the product of a motion* (Equations 22c,d) *and an axial inversion or of a motion* (Equations 22a,b) *and a dilatation.* The first case occurs when Equation 1 takes a pair of lines differing only in direction, z and $z_1 = -1/\bar{z}$, into lines differing only in direction,

$$z' = \frac{az+b}{cz+d} \qquad \text{and} \qquad z_1' = \frac{az_1+b}{cz_1+d} = \frac{-a+b\bar{z}}{-c+d\bar{z}} = -\frac{1}{\bar{z}'}$$

Thus we arrive at the equation

$$\frac{az+b}{cz+d} = -\frac{-\bar{c}+\bar{d}z}{-\bar{a}+\bar{b}z}$$

which can be rewritten in the form

$$(az+b)(-\bar{a}+\bar{b}z) + (cz+d)(-\bar{c}+\bar{d}z) = 0$$

or the form

$$Az^2 + 2Bz - \bar{A} = 0 \tag{30}$$

where

$$A = a\bar{b} + c\bar{d} \quad \text{and} \quad B = \tfrac{1}{2}(-a\bar{a} + b\bar{b} - c\bar{c} + d\bar{d}), \qquad B \text{ real} \tag{30a}$$

We must find out when Equation 30 has a solution. We suppose at first that $|A| \neq 0$, that is, that $A = t(1+\varepsilon\alpha)$, where $t \neq 0$. Putting $z = r(1+\varepsilon\phi)$, we obtain

$$tr^2[1+\varepsilon(\alpha+2\varphi)] + 2Br(1+\varepsilon\varphi) - t(1-\varepsilon\alpha) = 0$$

or

$$[tr^2 + 2Br - t] + [(tr^2 + 2Br - t)\varphi + (tr^2 + t)(\alpha + \varphi)]\varepsilon = 0$$

It is clear from this that the number $z = r(1 + \varepsilon\varphi)$ is a solution of Equation 30, where

$$\varphi = -\alpha, \qquad tr^2 + 2Br - t = 0, \qquad r = [-B + (B^2 + t^2)^{1/2}]/t$$

Further, if $|A| = 0$ but $A \neq 0$ (that is, if $A = \varepsilon\alpha$, where $\alpha \neq 0$), then by putting $z = \varepsilon\varphi$ we obtain from Equation 30:

$$2B \cdot \varepsilon\varphi + \varepsilon\alpha = 0$$

Thus, if $B \neq 0$, then the solution of Equation 30 has the form

$$z = -\frac{\alpha}{2B}\varepsilon$$

Finally, if $A = 0$, then the root of Equation 30 is

$$z = 0$$

However, if

$$A = \varepsilon\alpha, \qquad \alpha \neq 0, \quad B = 0$$

then, putting $z = r(1 + \varepsilon\varphi)$, we obtain

$$\varepsilon\alpha \cdot r^2 + \varepsilon\alpha = 0$$

which is impossible for any r. Putting $|z| = 0$ and $z = \varepsilon\varphi$, we arrive at the equation $\varepsilon\alpha = 0$, which also cannot hold. Thus we conclude that *the axial homography* (Equation 1) *cannot be expressed as a product of a motion and an axial inversion, if and only if*:

$$a\bar{b} + c\bar{d} \neq 0, \qquad |a\bar{b} + c\bar{d}| = 0, \qquad a\bar{a} + c\bar{c} = b\bar{b} + d\bar{d} \quad (31)$$

that is, if in Equation 30 $A \neq 0$, $|A| = 0$, and $B = 0$.

We now find when the transformation (Equation 1) can be expressed as the product of a motion and a dilatation. Let it be the product of the transformations (Equations 22a and 28); then we have

$$z' = \frac{az + b}{cz + d} = \frac{z_1 + \varepsilon(t/2)}{-\varepsilon(t/2)z_1 + 1}$$

where $$z_1 = \frac{pz + q}{-\bar{q}z + \bar{p}}, \qquad |\varDelta| = |p\bar{p} + q\bar{q}| \neq 0$$

Hence we obtain

$$z' = \frac{az + b}{cz + d} = \frac{[p - \varepsilon(t/2)\bar{q}]z + [q + \varepsilon(t/2)\bar{p}]}{[-\varepsilon(t/2)p - \bar{q}]z + [-\varepsilon(t/2)q + \bar{p}]} \tag{32}$$

Therefore we can put

$$\begin{aligned} a &= p - \varepsilon\frac{t}{2}\bar{q}, \qquad & b &= q + \varepsilon\frac{t}{2}\bar{p}, \\ c &= -\varepsilon\frac{t}{2}p - \bar{q}, \qquad & d &= -\varepsilon\frac{t}{2}q + \bar{p} \end{aligned} \tag{32a}$$

From Equations 32a it follows that

$$a\bar{b} + c\bar{d} = -\varepsilon t(p^2 + \bar{q}^2) = -\varepsilon t|\, p^2 + \bar{q}^2\,| = -\varepsilon t|\,\varDelta\,| \neq 0$$

$$|\, a\bar{b} + c\bar{d}\,| = 0$$

$$a\bar{a} + c\bar{c} = b\bar{b} + d\bar{d} = p\bar{p} + q\bar{q} = \varDelta$$

Therefore, *if the transformation* (Equation 1) *can be expressed as the product of a motion* (Equation 22a) *and a dilatation* (Equation 28), *then the conditions* (Equations 31) *are automatically satisfied.*

Conversely, let the circular transformation (Equation 1) be such that the conditions (Equations 31) are satisfied; let us assume also that $|\,a\,|$ and $|\,d\,|$ have the same sign. We shall show that our transformation can be expressed in the form of Equation 32, that is, that it consists of the product of a motion (Equation 22a) and a dilatation (Equation 28). First of all, it follows from Equations 31 that

$$|\, a\bar{b} + c\bar{d}\,| = |\,a\,| \cdot |\,b\,| + |\,c\,| \cdot |\,d\,| = 0,$$

$$|\,a\,| : |\,c\,| = -|\,d\,| : |\,b\,|, \qquad |\,a\,|^2 + |\,c\,|^2 = |\,d\,|^2 + |\,b\,|^2$$

Since we have assumed that $|\,a\,|$ and $|\,d\,|$ have the same sign, it follows that

$$|\,a\,| = |\,d\,| \qquad \text{and} \qquad |\,c\,| = -|\,b\,|$$

By virtue of this and Equation 32a we can determine p, q, and t:

$$p = \frac{a + \bar{d}}{2}, \qquad q = \frac{b - \bar{c}}{2}, \qquad 2\,\frac{\bar{d} - a}{\bar{b} - c} = 2\,\frac{b + \bar{c}}{\bar{a} + d} = \varepsilon t \tag{33}$$

Thus, *if the conditions* (Equations 31) *are satisfied and* $|\,a\,|$ *and* $|\,d\,|$ *have the same sign, then an axial circular transformation* (Equation 1) *can be expressed as the product of a motion* (Equation 22a) *and a dilatation* (Equation 28). In the same way, it may be shown that *the conditions* (Equations 31) *and the following equations* (we arrive at these equations by assuming that $|\,a\,|$ and $|\,d\,|$ have different signs),

$$|\,a\,| = -|\,d\,|, \qquad |\,c\,| = |\,b\,|$$

are necessary and sufficient conditions that the transformation (Equation 1) *may be expressed as the product of a motion* (Equation 22b) *and a dilatation* (Equation 28).

We have been speaking all the time of a *direct* axial circular transformation only for definiteness. All the reasoning applied to this can be carried over almost without change to an *opposite* axial circular transformation; here the conditions that a transformation (Equation 1a) can be expressed as the product of a motion (Equations 22c,d) and a dilatation (Equation 28) have the same form (Equations 31). Thus we can assert that *every axial circular transformation, such that the conditions are not satisfied, can be expressed as the product of a motion and an axial inversion, but if the conditions hold, every transformation can be expressed as the product of a motion and a dilatation.*

We have already observed (see p. 158) that *axial circular transformations form a group*. This enables us to call the study of the properties of figures, which are preserved by all such transformations, a special branch of geometry; this branch may be given the title of **axial circular geometry** [¶L]. In the following section we shall introduce some examples of theorems which refer to axial circular geometry.

*§16. Applications and Examples

The condition that four given (oriented) lines z_1, z_2, z_3, and z_4 of the plane can be taken by an axial circular transformation into four other lines w_1, w_2, w_3, and w_4 (the condition of **equality**

of these tetrads of lines in the sense of axial circular geometry) is, as we know, that *the tangential distance of the circles* $[z_1z_2z_3]$ *and* $[z_1z_2z_4]$ *should be equal to the tangential distance of the circles* $[w_1w_2w_3]$ *and* $[w_1w_2w_4]$ *and that the cross-ratio of the angles between the lines* z_1, z_2, z_3, *and* z_4 *should be equal to the cross-ratio of the angles between the lines* w_1, w_2, w_3, *and* w_4 (see p. 163):

$$([z_1z_2z_3], [z_1z_2z_4]) = ([w_1w_2w_3], [w_1w_2w_4])$$
$$\hat{W}(z_1, z_2, z_3, z_4) = \hat{W}(w_1, w_2, w_3, w_4)$$

But since

$$\{[z_1z_2z_3]z_1[z_1z_2z_4]\} = \tfrac{1}{2}(\{[z_1z_3], [z_2z_3]\} + \{[z_2z_4], [z_1z_4]\} - \{[z_4z_1], [z_3z_1]\} - \{[z_3z_2], [z_4z_2]\})$$
$$\hat{W}(z_1, z_2, z_3, z_4) = \frac{\sin(\frac{1}{2}\angle\{z_3, z_1\}) \cdot \sin(\frac{1}{2}\angle\{z_4, z_2\})}{\sin(\frac{1}{2}\angle\{z_3, z_2\}) \cdot \sin(\frac{1}{2}\angle\{z_4, z_1\})}$$

we obtain, in particular, the result: *a convex quadrilateral*[86] $\overline{z_1z_2z_3z_4}$ *can be taken by an axial circular transformation into another convex quadrilateral* $\overline{w_1w_2w_3w_4}$, *if and only if the difference*

$$(\{A, B\} + \{C, D\}) - (\{D, A\} + \{B, C\})$$

where

$$A \equiv [z_1z_3], \qquad B \equiv [z_2z_3], \qquad C \equiv [z_2z_4], \qquad D \equiv [z_1z_4]$$

between the sums of opposite sides of the first quadrilateral is equal to the difference of the sums of opposite sides of the second quadrilateral, and the ratio

$$\frac{\sin(\frac{1}{2}\angle\{z_3, z_1\}) \cdot \sin(\frac{1}{2}\angle\{z_4, z_2\})}{\sin(\frac{1}{2}\angle\{z_3, z_2\}) \cdot \sin(\frac{1}{2}\angle\{z_4, z_1\})}$$

of the products of the sines of half the opposite angles of the first quadrilateral is equal to the ratio of the products of the sines of half

---

[86] Here a quadrilateral $\overline{z_1z_3z_2z_4}$ is called **convex** if it lies entirely on one side (left or right) of each of its sides z_1, z_2, z_3, and z_4. (We note that the angles $\angle\{z_1, z_3\}$, $\angle\{z_1, z_4\}$, etc., coincide with the *exterior* angles of the quadrilateral, as usually understood.)

the opposite angles of the second quadrilateral. Hence it follows that *every convex quadrilateral can be taken by an axial circular transformation into a parallelogram.* If the original quadrilateral can be circumscribed about a circle, then this parallelogram is a rhombus; if the products of the sines of half the opposite angles of the quadrilateral are equal, then the parallelogram is a rectangle; finally, if both these conditions are satisfied at the same time, then the quadrilateral can be taken by an axial circular transformation into a square. Thus, a **harmonic quadrilateral** (see the end of Section 10) can be characterized as one *which can can be taken into a square by an axial circular transformation* (one which is equal to a square in the sense of axial circular geometry).

From what has been said it follows that all *axial circular properties* of a harmonic quadrilateral (those of its properties which are preserved by axial circular transformations) coincide with the corresponding properties of a square; the axial circular properties of a quadrilateral $\overline{z_1 z_2 z_3 z_4}$ which can be circumscribed about a circle coincide with the properties of a rhombus, and so on. This can be used to deduce many properties of quadrilaterals (cf. pp. 146 ff.); however, we shall leave the reader to do this for himself.

Here is another example of a theorem which can be proved by using axial circular transformations: *if four* (oriented) *circles* S_1, S_2, S_3, *and* S_4 *are such that* S_1 *and* S_3 *touch* S_2 *and* S_4, *then the common tangents* w_1, w_2, w_3, *and* w_4 *to the circles* S_1 *and* S_2, S_2 *and* S_3, S_3 *and* S_4, *and* S_4 *and* S_1, *drawn at the points of contact, touch one circle* Σ; see Figure 68a and also Figure 62a and text with it. To prove this we take the circle S_1 into a point S_1'; let the circles S_2, S_3, and S_4 go into circles S_2', S_3', and S_4' and let the lines w_1, w_2, w_3, and w_4 go into lines w_1', w_2', w_3', and w_4'. The points of contact of the circles S_2' and S_3' and of S_3' and S_4' will be denoted by A and B, and the points of intersection of the lines w_2' and w_3' with the lines w_1' and w_4' and with each other will be denoted by M, N, and P; see Figure 68b. By a well-known property of tangents to circles we have

$$\{S_1', M\} = \{M, A\}, \qquad \{B, N\} = \{N, S_1'\}, \qquad \{B, P\} = \{P, A\}$$

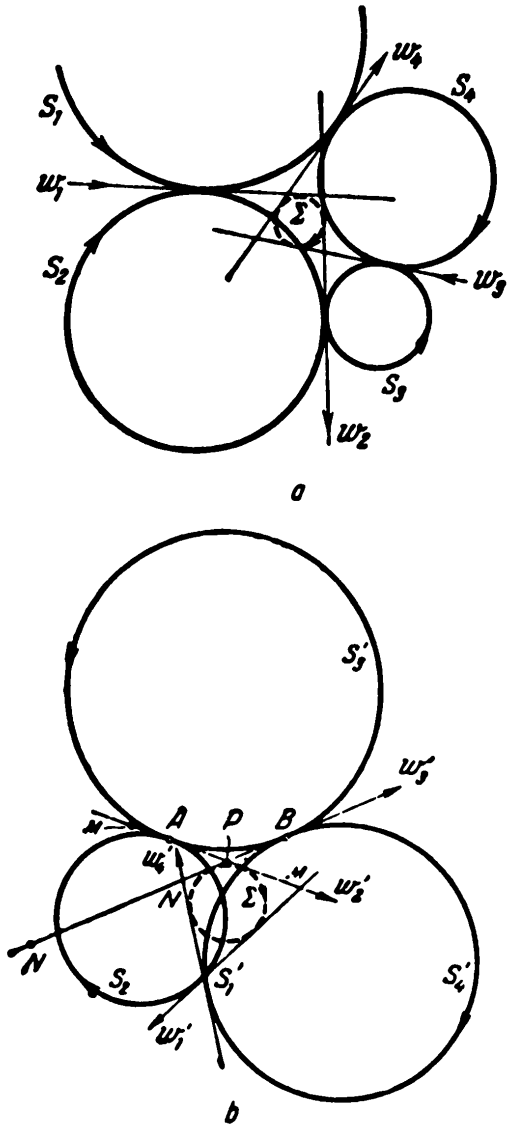


FIG. 68

Adding the first two of these equations and subtracting the third from their sum, we obtain

$$\{S_1', M\} + \{P, N\} = \{M, P\} + \{N, S_1'\}$$

whence it follows that a circle Σ' can be inscribed in the quadrilateral $S_1'MPN$ (that is, there exists a circle Σ' which touches the four oriented lines w_1', w_2', w_3', and w_4'; cf. p. 92).

We now turn to the question of the connection between axial inversion and the idea of the power of an (oriented) line with

respect to an (oriented) circle, a connection which enables us to give a "geometrized" description of the transformation of inversion. The **power of a circle** S, given by

$$Az\bar{z} + Bz - \bar{B}\bar{z} + C = 0, \qquad A, C \text{ purely imaginary}$$

(more precisely, *the power of the line o with respect to the circle S*) has been defined as the ratio $C/A = k$ of the first and last coefficients of the equation.

This power is a real number (or the "number" ∞; the power k is equal to ∞ if $A = 0$). Just as on pp. 155–156, it may be shown that the axial inversion of power k, Equation 25b,

$$z' = \frac{k}{\bar{z}}$$

takes S into itself. Hence it follows that if z_0 and z_0' are two (oriented) tangents to the circle S which meet at a point M of the axis o, then

$$z_0' = \frac{k}{\bar{z}_0}, \qquad \bar{z}_0 \cdot z_0' = k$$

(because these lines correspond to each other under the inversion (Equation 25b), and so

$$\tan\frac{\angle\{o, z_0\}}{2} \cdot \tan\frac{\angle\{o, z_0'\}}{2} = k \tag{34}$$

The tangents t_0 and t_0' to the circle S, such that

$$\bar{t}_0 \cdot t_0' = k \qquad \text{or} \qquad \frac{\{o, t_0\}}{\{o, t_0'\}} = -k \tag{34a}$$

are parallel and antiparallel to o; see Figure 69. Thus, we again arrive at the definition, introduced in Section 10, of the power k of the circle S as the product in Equation 34, where z_0 and z_0' are the two (oriented) tangents of S drawn to this circle from any arbitrary point M of the axis o.

The set of all circles (Equation 19) which go into themselves under the inversion (Equation 25b) with fixed power k, that is, the set of circles (Equation 19a) having the same power k,

$$Az\bar{z} + Bz - \bar{B}\bar{z} + Ak = 0$$

is called a **net of circles**; the number k is called the **power of**

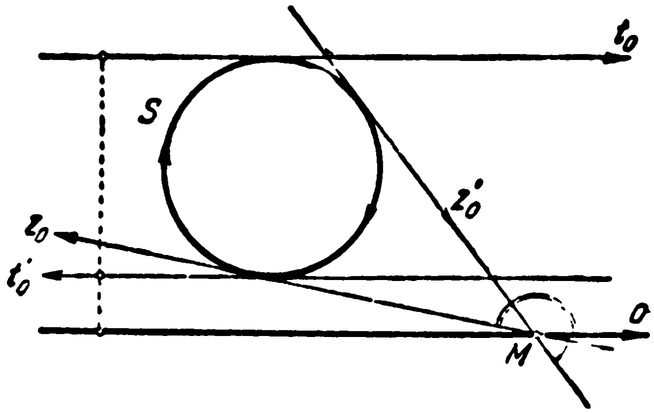


FIG. 69

the net, and the line o its **axis**. Since the power of S (the power of o with respect to S) is equal to $(r-d)/(r+d)$, where r is the (positive, zero, or negative) radius of S and d is the (positive, zero, or negative) distance of the center of S from o (see Section 10, especially Equation 45), *a net consists of all* (oriented) *circles of the plane for which*

$$\frac{r-d}{r+d}=k, \qquad \text{or} \quad \frac{r}{d}=\frac{1+k}{1-k}$$

In particular, if k is positive, the net consists of the set of circles which intersect o at a constant angle φ, such that $\tan^2 \varphi/2 = k$ (since the power of a circle S which intersects o at an angle φ is equal to $\tan^2 \varphi/2$);if $k=0$, the net consists of all circles which touch o; if $k=\infty$, the net consists of all circles which antitouch o; if k is negative, the net consists of all circles which subtend a constant angle ψ, such that $-\tan^2 \psi/4 = k$, at the foot of the perpendicular dropped from the center of the circle to o. See Figure 70a-d, and cf. p. 101.

We return now to the axial inversion with axis o and power k. We choose any circle Σ of the net with axis o and power k; we call this circle a **director circle** of our axial inversion (thus, an axial inversion has infinitely many director circles). From the definition of axial inversion it follows that, if z and z' are lines which correspond to each other under this inversion, then

$$\tan\frac{\angle\{o, z\}}{2}\cdot\tan\frac{\angle\{o, z'\}}{2}=k$$

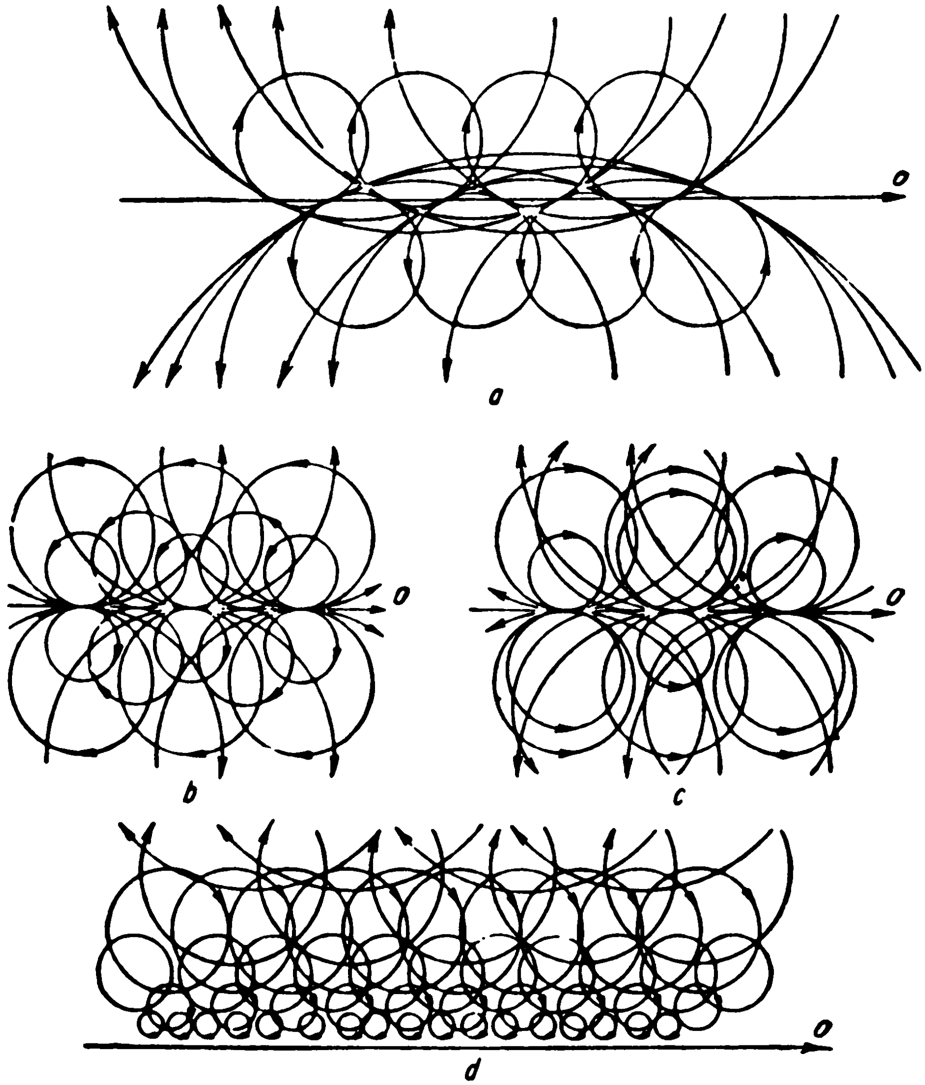


FIG. 70

or, if z is parallel and z' antiparallel to o,

$$\frac{\{o, z\}}{\{o, z\ \}} = -k$$

On the other hand, for the tangents z_0 and z_0' to the circle Σ which meet in the point M_0 of the axis o and for the tangents t_0 and t_0' which do not intersect o, Equations 34 and 34a hold. Hence it follows that *to construct the transform z' of the line z under the inversion* (Equation 25b) *it is sufficient to draw the tangent z_0 to the circle Σ parallel to z, the tangent z_0' to the same circle meeting o in the same point as z_0, and the line z' parallel to z_0' and meeting o*

in the same point as z (Figure 71); or, if z does not meet o, *to draw the tangent* t_0 *to the circle* Σ *parallel to* z, *the tangent* t_0' *to the same circle, which also does not meet* o, *and the line* z' *parallel to* t_0' *such that* $\{o, z\}/\{o, z'\} = \{o, t_0\}/\{o, t_0'\}$. This description of the construction of the line z' from the line z can be taken as the *definition of axial inversion* (given the axis o and the director circle Σ) [¶M].

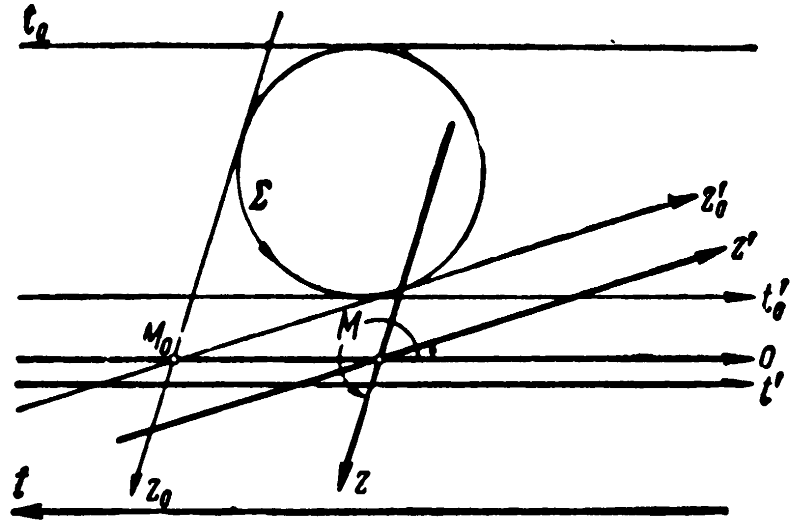


FIG. 71

The geometrical description of axial inversion and the theorem that every axial circular transformation may be represented as the product of a motion and an axial inversion or of a motion and a dilatation (see Section 15) enable us to give a sufficiently graphical geometrical description of any axial circular transformation. In particular, multiplication of a dual number z by a fixed number p

$$z' = pz$$

where p is real (Equations 56) may be represented as the product of the reorientation (Equation 21c) and the axial inversion (Equation 25b) of power $k = -p$:

$$z_1 = -\frac{1}{\bar{z}}, \qquad z' = -\frac{p}{\bar{z}_1}$$

If $p = r(1 + \varepsilon\varphi)$, then the transformation (Equation 5b) reduces to the translation $z_1 = (1 + \varepsilon\varphi)z$ in the direction of the axis o

through a distance φ, the reorientation $z_2 = -1/\bar{z}_1$, and the axial inversion $z' = -r/\bar{z}_2$ of power $-r$:

$$z' = -r/\bar{z}_2 = rz_1 = r(1 + \varepsilon\varphi)z$$

which is the *geometrical description of the multiplication of dual numbers*; see Figure 72a,b, in which $\tan^2 \alpha/2 = -r$ and

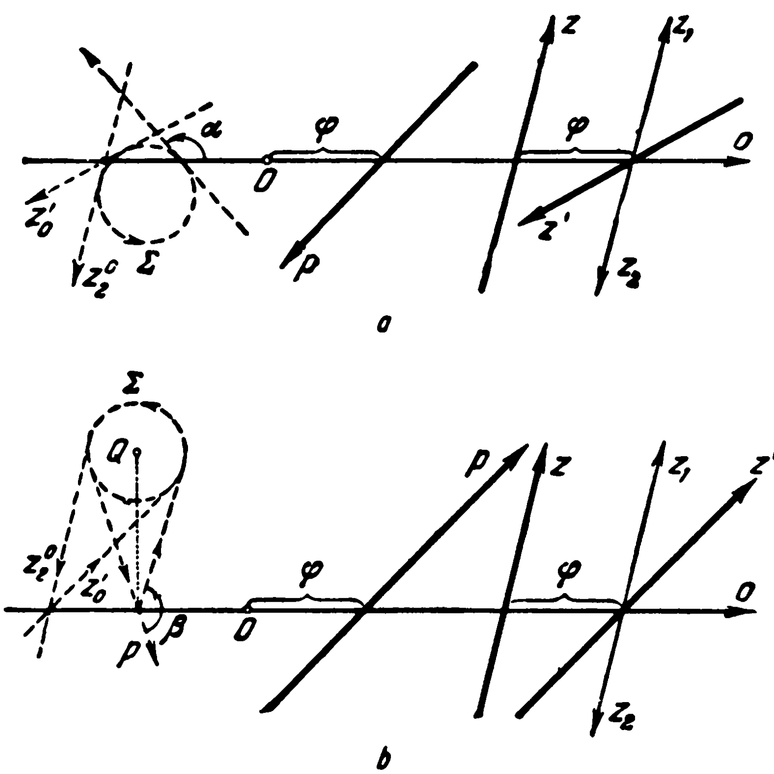


FIG. 72

$\tan^2 \beta/4 = r$. In the same way we may describe the transformation $z' = z + q$; that is, we may give the geometrical interpretation of the addition of two dual numbers. We shall leave the reader to do this.

**§17. Circular Transformations of the Lobachevskii Plane

The geometrical interpretation of ordinary complex numbers, explained in Section 11, enables us to consider the direct and

opposite linear-fractional transformations (Equations 1 and 1a) as transformations of the set of oriented points of the Lobachevskii plane. From the property of the invariance of the cross-ratio $W(z_1, z_2, z_3, z_4) = (z_1 - z_3)/(z_2 - z_3) : (z_1 - z_4)/(z_2 - z_4)$, given at the beginning of Section 13, and the condition, mentioned in Section 11, for four (oriented) points to lie on one (oriented) cycle it follows that these transformations *take every* (oriented) *cycle of the Lobachevskii plane into a cycle*, a result[87] which enables us to call Equations 1 and 1a *direct* and *opposite* **circular transformations**, or **Möbius transformations, of the Lobachevskii plane**. As in Section 13, it may be shown that there exists a unique direct (and a unique opposite) circular transformation which takes three given points z_1, z_2, and z_3 into three other known points w_1, w_2, and w_3. However, four (oriented) points z_1, z_2, z_3, and z_4 of the Lobachevskii plane can be taken into four known points w_1, w_2, w_3, and w_4, if and only if (Equation 9 or 9a):

$$W(w_1, w_2, w_3, w_4) = W(z_1, z_2, z_3, z_4)$$

or

$$W(w_1, w_2, w_3, w_4) = \overline{W(z_1, z_2, z_3, z_4)}$$

We now explain the geometrical significance of the cross-ratio of four points. It is not difficult to see that

$$\arg W(z_1, z_2, z_3, z_4) = \angle\{[z_1 z_2 z_4] z_1 [z_1 z_2 z_3]\}$$

where $\angle\{[z_1 z_2 z_4] z_1 [z_1 z_2 z_3]\}$ is the (oriented) **angle** which the cycles $[z_1 z_2 z_4]$ and $[z_1 z_2 z_3]$ passing through the points z_1, z_2, z_4 and z_1, z_2, z_3, respectively, form at the point z_1.[88] Further, we may show that

$$|W(z_1, z_2, z_3, z_4)| = \frac{\sinh(\frac{1}{2}\{z_3, z_1\})}{\sinh(\frac{1}{2}\{z_3, z_2\})} : \frac{\sinh(\frac{1}{2}\{z_4, z_1\})}{\sinh(\frac{1}{2}\{z_4, z_2\})}$$

[87] This, however, follows immediately from the corresponding property of circular transformations of the Euclidean plane and the fact that cycles of Lobachevskii geometry are represented by circles of the Euclidean plane.

[88] In other words, $\angle\{[z_1 z_2 z_4] z_1 [z_1 z_2 z_3]\}$ is the (oriented) angle between the *tangents* to the cycles $[z_1 z_2 z_4]$ and $[z_1 z_2 z_3]$ at the point z_1. The

where, for example, $\{z_3, z_1\}$ is the *distance* (which may be imaginary: see Section 11) between the oriented points z_3 and z_1. It is advisable to denote the expression on the right-hand side of this formula by a special symbol, say $W^*(z_1, z_2, z_3, z_4)$; we shall call it the **cross-ratio of the distances between the points** z_1, z_2, z_3 and z_4 (cf. p. 138). Thus it follows from the property of the invariance of the cross-ratio of four points that *circular transformations of the Lobachevskii plane preserve angles between* (oriented) *cycles*[89] *and cross-ratios of distances between points,* which enables us to use these transformations in many problems.

We now go into the question of the geometrical description of all circular transformations of the Lobachevskii plane. We note first of all that *every linear-fractional transformation* (Equation 1 or 1a) *of the Lobachevskii plane which takes the absolute* $z\bar{z} = 1$ *into itself is a motion,*[90] where among the motions we include the reorientation of points $z' = 1/\bar{z}$:

$$z' = \frac{pz + q}{\bar{q}z + \bar{p}} \quad \text{or} \quad z' = \frac{p\bar{z} + q}{\bar{q}\bar{z} + \bar{p}}, \qquad \Delta = \begin{vmatrix} p & q \\ \bar{q} & \bar{p} \end{vmatrix} \neq 0 \quad (35)$$

Hence it follows that any two transformations which take *the same* cycle S into the absolute Σ differ only by a motion; that is, each of these two transformations consists of the product of the other and a motion. Therefore, the classification of all circular transformations of the Lobachevskii plane reduces to the determination

---

proof of this statement follows immediately from what was said in Section 13 (pp. 136–137) about the geometrical meaning of arg $W(z_1, z_2, z_3, z_4)$, where z_1, z_2, z_3, and z_4 are four arbitrary points of the *Euclidean* plane, and from the facts that cycles of the Lobachevskii plane are represented by Euclidean circles and that the angle between intersecting cycles is equal to the angle between the circles which represent them.

[89] Hence it follows that every circular transformation of the Lobachevskii plane preserves angles between arbitrary curves; that is, it is a **conformal transformation** (which, moreover, follows immediately from the facts that circular transformations of the Euclidean plane are conformal and that the non-Euclidean angle between arbitrary curves is equal to the ordinary angle between their representations on the Poincaré model).

[90] See footnote 45.

of certain "standard" transformations, which take a circle, a horocycle, or an equidistant curve (in particular, a line) into Σ.

An example of a circular transformation which takes a *circle* S with center O and radius ρ (the equation of the circle being $z\bar{z} = k^2 = \tanh^2 \rho/2$) into the absolute Σ is given by an **inversion of the first kind**:

$$z' = \frac{k}{\bar{z}} \tag{36}$$

This transformation takes each diameter of the circle S into itself. It takes each interior point of the circle S, oriented in the same way as O, into a point z' of the ray Oz, oriented in the opposite way to O, such that the distances r and r' of these points from O are connected by the following relation (it also takes z' into z):

$$\tanh \frac{r}{2} : \tanh \frac{r'}{2} = k = \tanh \frac{\rho}{2}$$

Thus, each radius OM of the circle S (M is a point of the circle) is "stretched" into the whole ray OM. The transformation, Equation 36, takes an exterior point z of the circle, oriented in the same way as O, into a point z' of the ray Oz, such that the distances r and r' of the points z and z' from O are connected by the following relation (z and z' are oriented in the same way):

$$\tanh \frac{r}{2} \cdot \tanh \frac{r'}{2} = k = \tanh \frac{\rho}{2}$$

Thus, the segment of the ray OM exterior to the circle goes into itself, where the point M goes into the point at infinity and the point at infinity goes into the point M.

Equation 36 may be described differently. It is not difficult to see that this transformation takes into itself the circle S_1 with equation

$$z\bar{z} = k = \tanh^2 \frac{\rho_1}{2}$$

This is the circle whose radius ρ_1 is defined by the relation

$\tanh \rho/2 = \tanh^2 \rho_1/2$. Each point z goes into a point z' of the ray Oz such that

$$\tanh\left(\tfrac{1}{2}\{O, z\}\right) \cdot \tanh\left(\tfrac{1}{2}\{O, z'\}\right) = \tanh^2 \frac{\rho_1}{2}$$

In other words, each point z goes into a point z' such that all cycles of the non-Euclidean geometry of Lobachevskii perpendicular to S_1 which pass through z also pass through z' (cf. footnote 69); see Figure 73. The last property enables us to call

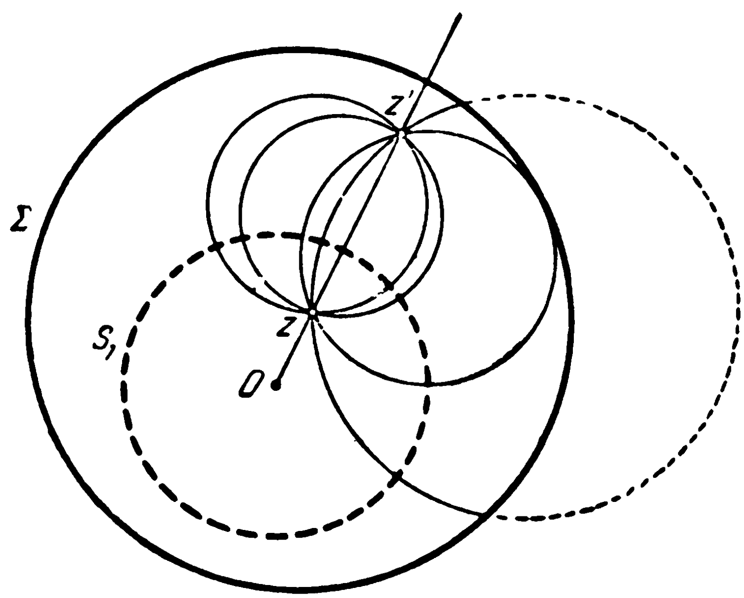


FIG. 73

the transformation given by Equation 36 the **symmetry with respect to the circle S_1**.

An example of a circular transformation which takes the *horocycle* S with equation $2z\bar{z} + iz - i\bar{z} = 0$ into the absolute is given by the **inversion of the second kind**:

$$z' = \frac{i\bar{z} + 1}{3\bar{z} + i} \tag{37}$$

This transformation takes each diameter of the horocycle into itself. Each point z interior to the horocycle, oriented in the same way as a point of the curve itself (that is, like the point O), goes into a point z', oriented in the opposite way to z, such that the

oriented distances d and d' of the points z and z' from the curve S, measured along the diameter $[zz']$, are connected by the relation

$$\cot\frac{\bar{d}}{2} - \cot\frac{\bar{d}'}{2} = 1$$

where $\bar{d}$ is the **angle of parallelism** corresponding to the segment d, and z' goes into z; thus, the ray of the diameter interior to S is "stretched" into the whole line. The transformation, Equation 37, takes a point z exterior to S, oriented in the same way as the points of S, into a point z' such that the distances d and d' of the points z and z' from S are connected by the following relation (z and z' are oriented in the same way):

$$\cot\frac{\bar{d}}{2} + \cot\frac{\bar{d}'}{2} = 1$$

Thus the ray of the diameter exterior to S goes into itself, where the point M of the curve S goes into the point at infinity and the point at infinity goes into M.

Equation 37 may be described differently. It is not difficult to see that this transformation takes into itself the horocycle S_1, defined by the equation

$$3z\bar{z} + iz - i\bar{z} - 1 = 0$$

Further, each point z goes into a point z', such that all cycles perpendicular to the curve S_1 which pass through the point z also pass through the point z'; see Figure 74. The transformation given by Equation 37 may be called the **symmetry with respect to the horocycle S_1**.

An example of a circular transformation which takes an *equidistant curve* S into the absolute, where the axis of S is o, the distance from points of S to the axis (the *width* of the equidistant curve) is h, and the equation of S is $z\bar{z}\sinh h - iz + i\bar{z} - \sinh h = 0$, is given by the **inversion of the third kind**:

$$z' = \frac{-(1-\alpha)i\bar{z} + (1+\alpha)}{(1+\alpha)\bar{z} - (1-\alpha)i}, \qquad \alpha = \tanh\frac{h}{2} \tag{38}$$

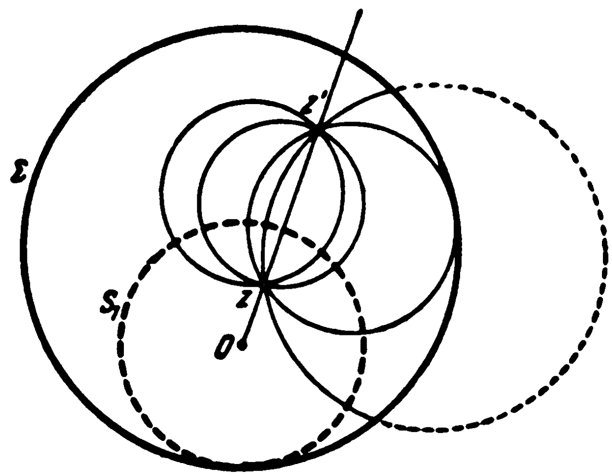


Fig. 74

This transformation takes each diameter of the equidistant curve into itself. The segment PM of the diameter between the axis and the equidistant curve, whose points are oriented in the same way as a point M of the equidistant curve, is "stretched" into the infinite ray bounded by a point M_1 of the diameter, such that $M_1P = h_1$ is the **complementary segment** of the segment $MP = h$. That is, the angles $\bar{h}$ and $\bar{h}_1$ are complementary $(\bar{h}_1 + \bar{h}_1 = \pi/2)$: if d and d' are the distances of a point z of the segment under consideration and its transform z' from the axis o of the equidistant curve, then

$$\bar{d} - \bar{d}' = \bar{h}$$

where, as before, $\bar{d}$ is the angle of parallelism, and the orientations of the points z and z' are opposite; conversely, z' goes into z. The ray of the diameter lying outside the equidistant curve and bounded by the point M goes into itself, where M goes into the point at infinity and the point at infinity goes into M: if d and d' are the distances of a point z of this ray and its transform z' from the axis o (z and z' are oriented in the same way as a point M of the equidistant curve), then

$$\bar{d} + \bar{d}' = \bar{h}$$

Finally, points of the segment PM_1 oriented in the opposite way to M go into points of a segment PM_1' equal to it and lying on the

other side of the axis: if d and d' are the distances from the axis o of points z and z' of the segments PM and $P'M$ which correspond to each other under the transformation given by Equation 38, then

$$\bar{d} + \bar{d}' = \bar{h}_1$$

Here d is regarded as positive, and the orientations of the points z and z' are opposite.

In particular, by taking the width h of the equidistant curve to be zero, we obtain the transformation

$$z' = \frac{-i\bar{z} + 1}{\bar{z} - i} \tag{39}$$

which takes the *line* o of the Lobachevskii plane into the absolute. This transformation takes each line perpendicular to o into itself. A point z of a ray of this line bounded by a point p of the line o, depending on its orientation, goes either into a point z' of this ray (if the non-Euclidean distances (p, z) and (p, z') are equal to d and d', then

$$\bar{d} + \bar{d}' = \pi/2\ ,$$

that is, d and d' are complementary segments), or it goes into a point z' of the second ray of this line (since in this case the segments $\overset{\frown}{pz}$ and $\overset{\frown}{pz'}$ are complementary); see Figure 75. The German geometer H. Liebmann first arrived at this transformation, by quite different considerations.[91]

The transformation (Equation 38) takes into itself the equidistant curve S_1 with equation

$$e^h z\bar{z} - iz + i\bar{z} - e^h = 0$$

which is the equidistant curve with axis o and width h_1 such that $\sinh h_1 = e^h$, where e is the base of natural logarithms, approximately equal to 2.7. It takes each point z of the plane into a point z' such that all cycles perpendicular to S_1 which pass through z also pass through z'; see Figure 76. This

[91] See H. Liebmann, *Nichteuklidische Geometrie*, §26 (Göschen, Leipzig), 1905. (We note that Liebmann does not introduce oriented points, which somewhat complicates the situation.)

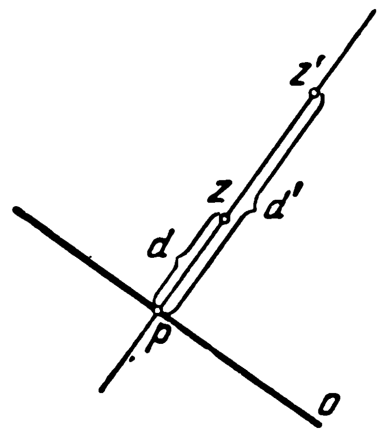


Fig. 75

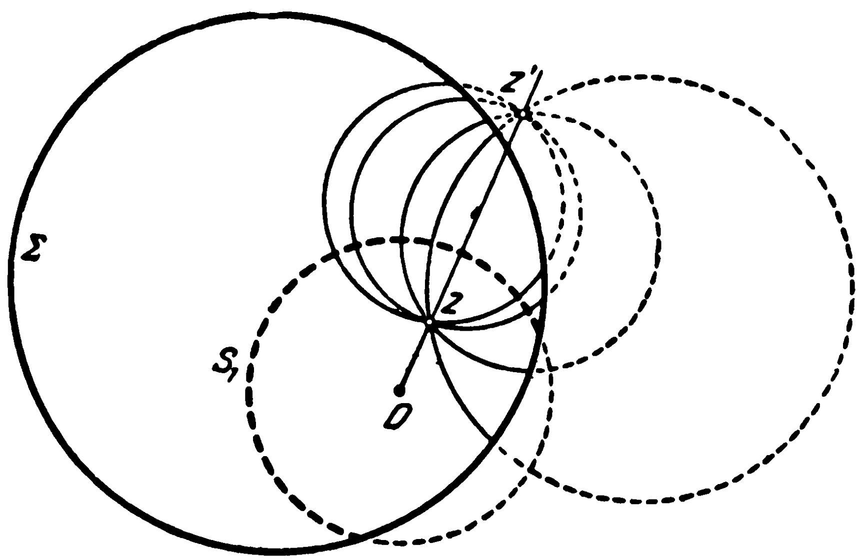


Fig. 76

transformation is naturally called the **symmetry with respect to the equidistant curve** S_1. In particular, Equation 39, called the **Liebmann transformation**, consists of the symmetry with respect to the equidistant curve of width h_1, where $\sinh h_1 = 1$.

Thus, the complete classification of circular transformations of the Lobachevskii plane gives the following theorem:

Every circular transformation of the Lobachevskii plane consists of either a motion, or a motion together with a symmetry with respect to some cycle S_1 (an inversion of the first, second, or third kind).

We note that every linear-fractional (circular) transformation in the plane of the complex variable provides us with a definite modification of the Poincaré model of Lobachevskii geometry. In fact, on the original Poincaré model, let a point A be represented by a point (complex number) z. We obtain the modification of the Poincaré model by agreeing to represent the point A of the Lobachevskii plane, not by the point z, but by the point (complex number) z' into which our transformation takes the point z. The properties of circular transformations enable us to state that on the modified Poincaré model obtained in this way cycles are represented by circles of the plane of the complex variable, and the non-Euclidean angle between cycles is equal to the (ordinary) angle between the circles which represent these cycles. In particular, the so-called **Cayley transformation**

$$z' = \frac{iz + i}{-z + 1} \tag{40}$$

which takes the circle $z\bar{z} = 1$ into the real axis $z - \bar{z} = 0$ maps the Poincaré unit circle model onto the Poincaré half-plane model (see pp. 116-117).

**§18. Axial Circular Transformations of the Lobachevskii Plane

We consider now the linear-fractional transformations (Equations 1 and 1a),

$$z' = \frac{az + b}{cz + d}, \qquad z' = \frac{a\bar{z} + b}{c\bar{z} + d}$$

where z and z' are *double* numbers; here, as before, we shall assume that the determinant $\Delta = \left| \begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right|$ is not a divisor of zero. By virtue of what was said in Section 12, these transformations can be interpreted as transformations in the set of **oriented lines** of the Lobachevskii plane. From the property of the invariance of the cross-ratio $W(z_1, z_2, z_3, z_4) = (z_1 - z_3)/(z_2 - z_3) : (z_1 - z_4)/(z_2 - z_4)$, which may be deduced exactly as above, and from the condition, mentioned in Section 12, for four oriented lines z_1, z_2, z_3, and z_4 to touch one cycle, it follows that *the transformations* (Equations 1 and 1a) *take each cycle of the*

Lobachevskii plane into a cycle; by virtue of this, such a transformation is naturally called a (*direct* or *opposite*) **axial circular transformation of the Lobachevskii plane**, and is sometimes also called a **Laguerre transformation of the Lobachevskii plane**.

The cross-ratio of four (oriented) lines z_1, z_2, z_3, and z_4 is a fundamental invariant of axial circular transformations; more precisely, if the transformation (Equation 1 or 1a) takes the lines z_1, z_2, z_3, and z_4 into the lines z_1', z_2', z_3', and z_4', then (Equations 9 and 9a):

$$W(z_1', z_2', z_3', z_4') = W(z_1, z_2, z_3, z_4)$$

or

$$W(z_1', z_2', z_3', z_4') = \overline{W(z_1, z_2, z_3, z_4)}$$

It is not difficult to explain the geometrical significance of the expression $W(z_1, z_2, z_3, z_4)$. The argument $\arg W$ of the cross-ratio is equal to the (oriented) tangential distance $\{[z_1z_2z_4]z_1[z_1z_2z_3]\}$, measured along the tangent z_1, between the cycles $[z_1z_2z_3]$ and $[z_1z_2z_4]$, determined by the triads of lines z_1, z_2, z_3 and z_1, z_2, z_4; the modulus $|W|$ of the cross-ratio is equal to the **cross-ratio of the angles between the lines**

$$\hat{W}(z_1, z_2, z_3, z_4) = \frac{\sin(\frac{1}{2}\angle\{z_3, z_1\})}{\sin(\frac{1}{2}\angle\{z_3, z_2\})} : \frac{\sin(\frac{1}{2}\angle\{z_4, z_1\})}{\sin(\frac{1}{2}\angle\{z_4, z_2\})}$$

determined exactly as in Section 15. From the property of the invariance of the cross-ratio it follows that *axial circular transformations of the Lobachevskii plane preserve tangential distances between cycles*[92] *and cross-ratios of angles between tetrads of lines*; this enables us to use axial circular transformations in many problems of non-Euclidean geometry.

[92] It follows from this that *axial circular transformations of the Lobachevskii plane preserve tangential distance between arbitrary curves*; that is, they are equidistantial transformations of the Lobachevskii plane (cf. footnote 80).

In conclusion we give a geometrical description of axial circular transformations of the Lobachevskii plane. It is not difficult to verify that *every axial circular transformation* (Equation 1 or 1a) *of the Lobachevskii plane which takes the absolute* $z\bar{z} = -1$ *into itself is a motion*:

$$z' = \frac{pz + q}{-\bar{q}z + \bar{p}}, \qquad z' = \frac{-pz + q}{\bar{q}z + \bar{p}}, \qquad z' = \frac{p\bar{z} + q}{-\bar{q}\bar{z} + \bar{p}},$$
$$\text{or} \quad z' = \frac{-p\bar{z} + q}{\bar{q}\bar{z} + \bar{p}}, \qquad p\bar{p} + q\bar{q} > 0 \tag{41}$$

Since, in addition, as we may easily see, *an axial circular transformation takes parallel pencils into parallel pencils*, the problem of the classification of all axial circular transformations of the Lobachevskii plane reduces to the determination of "standard" transformations, which take an equidistant curve, a horocycle, a circle (in particular, a point), or a pencil of equal inclination into the absolute Σ.

For what follows it will be convenient to introduce the following idea. With the (oriented) lines of the Lobachevskii plane plane perpendicular to a fixed axis l (lines of the orthogonal pencil with axis l) we associate the oriented points of intersection of these lines with l in such a way that to lines directed on the same side of l correspond similarly oriented points. By virtue of this correspondence, to each transformation in the set of oriented lines (axes) of an orthogonal pencil corresponds some transformation in the set of oriented points lying on a line l. We shall say that these transformations, both point and axial, are **compatible**.

An example of an axial circular transformation which takes an *equidistant curve* S with axis o and width h, given by the equation $z\bar{z} = -k^2 = -\tanh^2 h/2$, into the absolute Σ is the **axial inversion of the first kind**:

$$z' = -\frac{k}{\bar{z}} \tag{42}$$

Under this transformation each line z, which intersects the axis of the equidistant curve S in some point M, goes into the line z',

which intersects o in the same point M, and satisfies the condition

$$\tan \frac{\{o, z\}}{2} \cdot \tan \frac{\{o, z'\}}{2} = k$$

A line z parallel to o in one of the directions of this line goes into a line z' antiparallel to o in the opposite direction, and z' goes into z; if p and p' are the (oriented) distances of z and z' from the pole O of the system of polar coordinates, where O lies on o, then

$$\sinh p/\sinh p_1 = -k$$

Finally, each orthogonal pencil, constructed on the diameter l of the equidistant curve S, goes into itself, and the transformation of the lines of this pencil is compatible with the transformation to which the points of the diameter l of the circle $\bar{S}$ "inscribed" in the equidistant curve S (the radius of the circle $\bar{S}$ is equal to h, and its center coincides with the point P where l intersects o) are subjected under the point inversion of the first kind, which takes S into the absolute.

An example of an axial circular transformation which takes a *horocycle* S with equation $2z\bar{z} + ez - e\bar{z} = 0$ into the absolute Σ is the **axial inversion of the second kind**:

$$z' = \frac{e\bar{z} - 1}{3\bar{z} + e} \tag{43}$$

Under this transformation the pencil of diameters of the horocycle S goes into itself. The orthogonal pencil whose axis is the diameter l of the curve S also goes into itself, and the transformation of the lines of this pencil is compatible with the transformation of the points of the diameter l under the point inversion of the second kind (Equation 37).

An example of an axial circular transformation which takes the *circle* S, of radius r with center at the pole O of the coordinate system (the equation of this circle has the form $z\bar{z} \sinh r - ez +$ $+ e\bar{z} + \sinh r = 0$), into the absolute Σ is given by the **axial inversion of the third kind**:

$$z' = \frac{(1 - \alpha)e\bar{z} + (1 + \alpha)}{-(1 + \alpha)\bar{z} + (1 - \alpha)e}, \qquad \alpha = \tanh \frac{r}{2} \tag{44}$$

Under this transformation each orthogonal pencil, whose axis is a diameter l of the circle, goes into itself, and the transformation of the lines of this pencil is compatible with the transformation of the points of l under the point inversion of the first kind, which takes the equidistant curve $\bar{S}$ circumscribed about S into itself (the width of the equidistant curve $\bar{S}$ is equal to r, and its axis is perpendicular to l and meets l in the point O). In particular, putting the radius r of the circle equal to zero, we arrive at the transformation

$$z' = \frac{e\bar{z} + 1}{-\bar{z} + e} \tag{45}$$

which takes the *point* O into the absolute Σ. Under this transformation the line z goes into the line z' perpendicular to the line $OP \perp z$ and meeting it in the point P' such that the distances $OP = d$ and $OP' = d'$ from O to z and z' are complementary segments:

$$\bar{d} + \bar{d}' = \frac{\pi}{2}$$

See Figure 77.

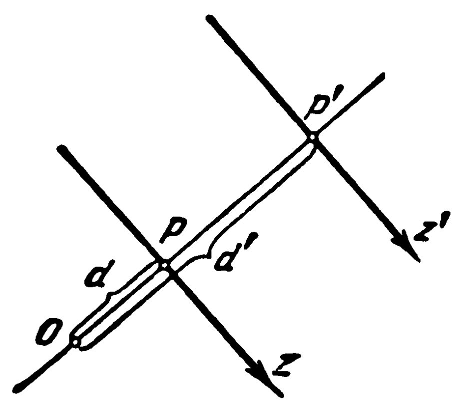


FIG. 77

We note that the transformations given by Equations 42, 43, and 44 leave fixed all the (oriented) lines which touch respectively the equidistant curve S_1 with equation $z\bar{z} = -k$, the horocycle S_1 with equation $3z\bar{z} + ez - e\bar{z} + 1 = 0$, or the circle S_1 with equation $az\bar{z} - ez + e\bar{z} + a = 0$, where a is a real number

connected with the radius r of the circle S by the relation $a = e^r$, e being the base of natural logarithms (not the double unit!). Further, all these transformations take each point z into a point z', which in turn goes into the point z. So these transformations can also be called **axial symmetries with respect to an equidistant curve, a horocycle, or a circle** S_1, respectively. In particular, Equation 45 represents the axial symmetry with respect to the circle S_1, whose radius r_1 is determined by the relation $\sinh r_1 = 1$.

We consider finally the following special axial circular transformation of the Lobachevskii plane:

$$z' = ez \tag{46}$$

This transformation takes each line z which intersects the axis o into a line z' ultraparallel to o, and z' into z; it takes a line z parallel to o into a line z' parallel to o (z and z' are parallel to o in different directions, and z' goes into z), and it similarly transforms lines antiparallel to o. A pencil of lines which intersects o at some point M goes into an orthogonal pencil, whose axis is the perpendicular l to the axis o erected at the point M; the **angle** $\varphi = \angle\{o, z\}$ is connected with the **distance** $d = \{o, z'\}$ from o of the transform z' of the line z by the relation

$$\begin{aligned} \tanh\frac{d}{2} &= \tan\frac{\varphi}{2}, &\quad \text{if } \left|\tan\frac{\varphi}{2}\right| \leqslant 1 \\ &= \cot\frac{\varphi}{2}, &\quad \text{if } \left|\tan\frac{\varphi}{2}\right| > 1 \end{aligned}$$

and if $|\tan\varphi/2| \leqslant 1$ the line z' is directed on the same side of l as o, but if $|\tan\varphi/2| > 1$ it is directed on the opposite side. The pencil of lines which intersect the polar axis o at some constant angle $\varphi \neq \pi/2$ goes into the equidistant curve with axis o; the orthogonal pencil with axis o goes into the absolute Σ; conversely, Σ goes into the orthogonal pencil with axis o. The transformation Equation 46 takes each equidistant curve, horocycle, or circle into a pencil of equal inclination, and conversely; in particular, it takes (ordinary) points into pencils of equal inclination. This transformation may be called a **reversion**.

It is now not difficult to find an example of a transformation which takes a pencil of equal inclination φ with axis o into Σ (a pencil of lines which intersect o at a constant angle φ, the equation of which has the form $z\bar{z} = k^2 = \tan^2 \varphi/2$). This example is given by the product of an axial inversion of the first kind (Equation 42) and the reversion (Equation 46):

$$z' = -\frac{k}{\bar{z}_1}, \qquad z_1 = ez$$

or (47)

$$z' = \frac{ke}{\bar{z}}$$

This transformation can be called the **axial inversion of the fourth kind**.

Finally we obtain the result that *each axial circular transformation of the Lobachevskii plane represents a motion, or a motion together with an axial symmetry with respect to some cycle* S_1 (axial inversion of the first, second, or third kind), *or a motion together with an axial symmetry with respect to an equidistant curve and a reversion* (that is, together with an axial inversion of the fourth kind).

In conclusion, we show that each axial circular transformation enables us to find a new mapping of the set of directed lines of the Lobachevskii plane onto the set of double numbers or, in other words, new "complex coordinates" (more precisely, double coordinates) of lines; for this it is sufficient to associate with a line, which previously had the "coordinate" z, the number $z' = (az + b)/(cz + d)$. In the new coordinate system the condition that four lines z_0, z_1, z_2, and z_3 should touch one cycle has the previous form $W(z_0, z_1, z_2, z_3) = \overline{W(z_0, z_1, z_2, z_3)}$; cycles are expressed by the familiar Equations 19; axial circular transformations have the forms given by Equation 1 and 1a; and so on. In particular, the transformation

$$z' = \frac{ez + 1}{z - e} \tag{48}$$

takes the complex coordinates (given by Equations 59, 59a, and 59b of Chapter II) of lines of the Lobachevskii plane into those complex coordinates considered at the end of Section 12 (pp. 126–129).

APPENDIX

Non-Euclidean Geometries in the Plane and Complex Numbers

A1. Non-Euclidean Geometries in the Plane

After the discovery in the first half of the 19th century by K. F. Gauss (1777–1855), J. Bolyai (1802–1860), and N. I. Lobachevskii (1792–1856) of a new geometrical system, different from the classical geometry of Euclid, mathematicians were forced to relinquish their traditional belief in the logical uniqueness of the geometry of Euclid which they knew from their schooldays; it was established that there exist *two* geometries of equal status, Euclidean and non-Euclidean. The title "non-Euclidean geometry" for this new system of geometry became fixed: today it is more often called the *non-Euclidean geometry of Lobachevskii** or *hyperbolic geometry*, but even today the expression "non-Euclidean geometry," without any further explanation, denotes as a rule the non-Euclidean geometry of Lobachevskii. Apparently only B. Riemann (1826–1866) gave due consideration to the fact that as well as the non-Euclidean geometry of Lobachevskii (hyperbolic geometry) we have another geometrical system, different from the classical geometry of Euclid and in fact known to mathematicians much earlier than the non-Euclidean geometry of Lobachevskii; this was the geometry on the surface of a sphere in ordinary Euclidean space, where the distance between points is defined as the length of the shortest

* Lobachevskii was the author of the first published paper on the new geometry.

path joining the points on the surface of the sphere, and the role of lines is played by the "shortest curves," or geodesic curves, of the spherical surface—the great circles of the sphere. Since any two great circles of a sphere intersect in two diametrically opposite points, and we are accustomed to the fact that two lines always intersect in not more than *one* point, it is convenient to modify spherical geometry by agreeing to call a "point" of this geometry any pair of diametrically opposite points. The set of such pairs—a pair being considered as one element of the geometry, one "point"—is called an *elliptic plane*; the corresponding geometrical system is called today the *non-Euclidean geometry of Riemann*, or *elliptic geometry*.

Thus geometers were forced to recognize the existence of not just two but *three* geometrical systems of equal status: Euclidean geometry and the two non-Euclidean geometries of Lobachevskii and Riemann. From this time it became sensible to use the expression "non-Euclidean geometries," where the plural originally meant that it was a question of two geometrical systems, those of Lobachevskii and Riemann. However, the progress of geometry did not stop there.

In 1870, the famous German mathematician F. Klein (1849–1925), relying on a somewhat earlier investigation by the Englishman A. Cayley (1821–1895), described a common way of establishing all the three geometries known at that time—Euclidean geometry and the non-Euclidean geometries of Lobachevskii and Riemann.[1] However, Klein's constructions turned out to be more general: they led to a *whole series* of similar geometrical systems, which today are often called *Cayley-Klein non-Euclidean geometries* (or *projective metrics*). Thus today the words "non-Euclidean geometries," used in the plural, mean not only the classical non-Euclidean geometries of Lobachevskii and Riemann, but many other geometrical systems.[2]

---

[1] See F. Klein, *Nichteuklidische Geometrie* (Chelsea, New York), 1960.

[2] Contemporary geometry includes, as well as the Cayley-Klein non-Euclidean geometries, many other systems different from Euclidean, which might perhaps be called "non Euclidean geometries," without the hyphen.

It is rather inconvenient that the classical geometry of Euclid is included among these non-Euclidean geometries—however, in other respects this terminology is convenient, since all the Cayley-Klein geometries have many common features, which compel us to look for a simple common name for them.[3]

Klein's investigation was somewhat revised in 1910 by the British mathematician D. M. Y. Sommerville (1874–1939), who established the existence in the plane of *nine* different projective metrics, or Cayley-Klein non-Euclidean geometries. It is not difficult to describe these nine geometries. On a line we have three geometries (or metrics) of this type: *hyperbolic* (or Lobachevskian), *parabolic* (or Euclidean), and *elliptic* (or Riemannian). These metrics can be written in a common form by using the idea of the "absolute": The absolute of hyperbolic geometry is a pair of points I and J of the (projective) line (Figure A1a)[4]; the absolute of elliptic geometry is a pair of conjugate complex points I and J of the (projective) line (Figure A1b)[4]; finally, the absolute of parabolic or Euclidean geometry consists of two coincident points $I \equiv J$ of the line (Figure A1c)[4], which can be taken as the point at infinity, and if we exclude this our line becomes an affine (or Euclidean) line.

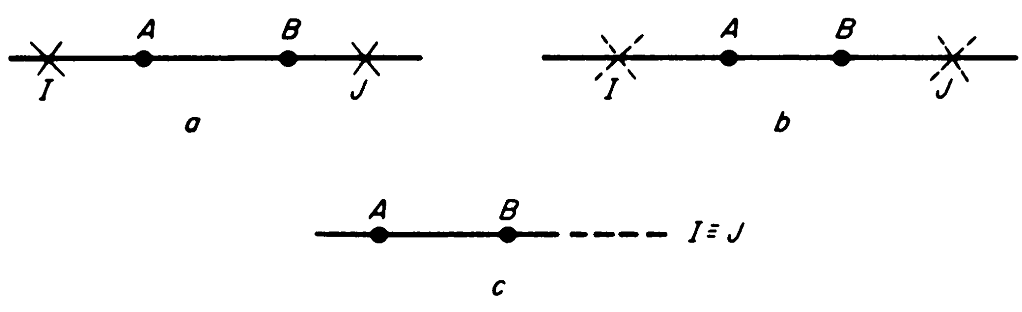


FIG. A1

[3] A somewhat different terminology is adopted in the article by I. M. Yaglom, B. A. Rozenfel'd, and E. U. Yasinskaya referred to on p. 14.

[4] In the projective coordinates (x_1, x_0) of a line, the equation of the absolute of a hyperbolic line, elliptic line, and Euclidean line may be written in the forms

$$x_1^2 - x_0^2 = 0, \quad x_1^2 + x_0^2 = 0, \quad x_0^2 = 0$$

The distance d_{AB} between the points A and B in the hyperbolic or elliptic case can be described by the same formula

$$d_{AB} = k \log W(A, B; I, J) \tag{A1}$$

where $W(A, B; I, J)$ is the *cross-ratio* of four points of the line, which plays such an important role in projective geometry, and k is some constant, the choice of which is connected with the choice of unit length of the corresponding metric; in order that the distance d_{AB} should be real in the elliptic case it is necessary to regard the constant k as purely imaginary. In the parabolic (Euclidean) case the cross-ratio $W(A, B; I, J)$, where I and J are *the same* point, is identically equal to 1; hence to define length we must impose some further conditions, which we shall not go into here.[5]

This description of the three non-Euclidean geometries on a line is rather too complicated, since it contains conjugate complex points, which are not intuitively obvious geometrical objects. We may, however, describe these three geometries while keeping within the bounds of elementary geometry. The distance d between the points A and B of a *hyperbolic* line is best described by formula A1:

$$d = \log W(A, B; I, J) = \log \frac{AI}{BI} : \frac{AJ}{BJ}$$

where we have put $k = 1$; in order that this distance should be real, we must agree to limit ourselves to points of the segment IJ (thus, this segment, excluding the end-points I and J, plays the role of the *complete* hyperbolic line). If A, C and B are three arbitrary points of a hyperbolic line, in that order (Figure A2a), then it is easily verified that

$$d_{AC} + d_{CB} = d_{AB}$$

Further, if the point B tends to one of the absolute points I or J,

---

[5] All projective metrics on a line can be written uniformly without resorting to the complex geometrical form; we shall not go into this here.

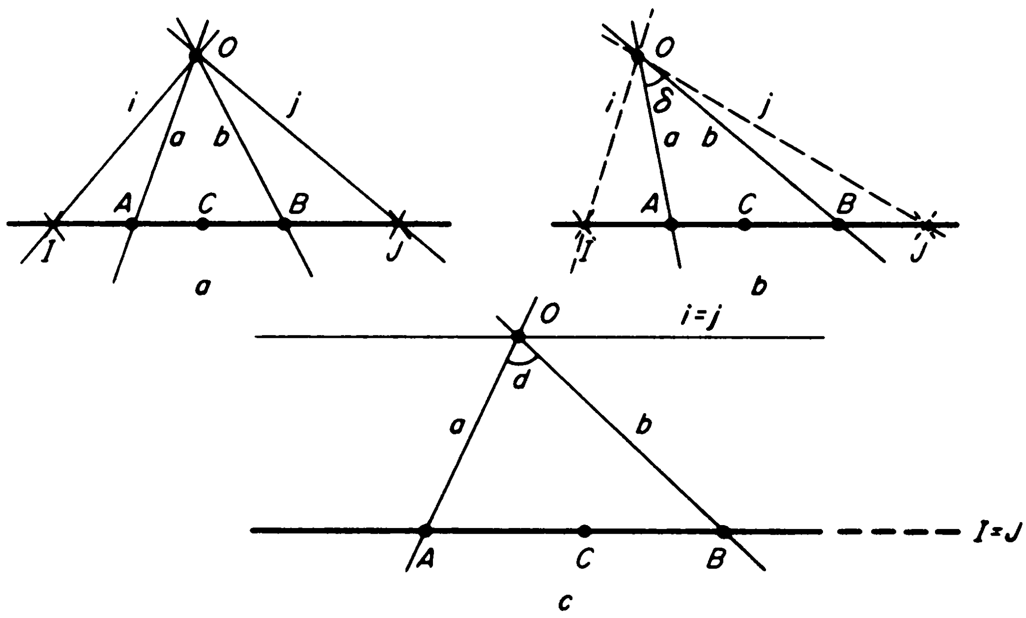


FIG. A2

then the distance d_{AB} tends to ∞ (that is, it increases without limit), so the length of the whole hyperbolic line is infinite.

It is easier to define a *Euclidean* metric on a line; strictly speaking, we need not define it again, since it is well known from school geometry. If EF is an arbitrary unit segment, fixed on the line, then d_{AB} is the (Euclidean; that is, described in detail in Euclid's *Elements*) ratio of segments

$$d = \frac{AB}{EF}$$

Here also, if A, C, and B are three points in that order on a Euclidean line (Figure A2c), then

$$d_{AC} + d_{CB} = d_{AB}$$

and if the point B tends to the point at infinity I, then d_{AB} tends to ∞. Finally, an *elliptic* metric on a line can be defined by putting

$$d_{AB} = \angle(a, b) = \angle AOB$$

where O is some point of the plane not lying on the line AB, and $\angle AOB$ is the ordinary (Euclidean) angle δ between the lines

$OA = a$ and $OB = b$ (Figure A2b). Here again, for three points A, C, and B of the line (for a suitable choice of all the distances d_{AB}, d_{AC}, and d_{CB}, which, just like the angle between rays of the Euclidean plane, are ambiguously defined),

$$d_{AC} + d_{CB} = d_{AB}$$

As for the whole elliptic line, its length should be regarded as coinciding with the value of a straight angle; that is, it should be counted as equal to π.

We now turn to the measure of angles between lines of a pencil. This measure can also be introduced in three different ways. The ordinary (that is, as in Euclidean geometry) metric of angles with a fixed vertex O (see Figure A2b) is called *elliptic.* We obtain the *parabolic* (or *Euclidean*) metric of angles by supposing that

$$\delta_{ab} = AB$$

where AB is the (Euclidean) length of the segment cut by the arms a and b of the angle on some line l not passing through the vertex of the angle (Figure A2c). Finally, the *hyperbolic* metric of angles is defined by the equation

$$\delta_{ab} = \log W(A, B; I, J)$$

where $\log W(A, B; I, J)$ is the hyperbolic length of the segment AB cut by the arms of the angle on a fixed line l not passing through the vertex O of the angle; here we suppose that a hyperbolic metric of distance is assigned to the line l (Figure A2a). In order that the hyperbolic magnitudes of angles should all be real, we must restrict ourselves to the measurement of angles between rays lying inside the fixed angle IOJ; the lines OI $(= i)$ and OJ $(= j)$ form the angular absolute, by means of which the magnitude of an angle can be described by the following formula[6]

[6] Also the elliptic measure of angles can be given by the formula related to (A1a)

$$\delta_{ab} = k \log W(a, b; i, j)$$

where $i \equiv OI$ and $j \equiv OJ$ are conjugate complex lines.

related to formula A1:

$$\delta_{ab} = \log W(a, b; i, j) \tag{A1a}$$

After this brief survey of the 1-dimensional Cayley-Klein non-Euclidean geometries, it is not difficult to examine the plane (or 2-dimensional) non-Euclidean geometries. According to the scheme of Sommerville, the nine projective metrics in the plane correspond to all possible combinations which can be formed from the three kinds of measures of distances and the three kinds of measures of angles. If we associate elliptic measure with the symbol -1, Euclidean (parabolic) measure with the symbol 0, and hyperbolic measure with the symbol $+1$, and agree that the symbol in the first position corresponds to the measure of distances and the symbol in the second position to the measure of angles, we arrive at the following nine geometries[7]:

1. *Elliptic geometry of Riemann* $[-1, -1]$ (elliptic measure of distances and angles). For the absolute of this geometry in the (projective) plane we may take any imaginary conic (Figure A3a).[8]

2. *Hyperbolic geometry of Lobachevskii* $[+1, -1]$ (hyperbolic measure of distances, elliptic measure of angles). For the absolute

---

[7] Similarly in 3-dimensional space we may define $3^3 = 27$ non-Euclidean geometries corresponding to all possible combinations of the three measures of distances between points of a line, angles between lines of a (plane) pencil, and angles between the planes of a pencil.

[8] In projective coordinates (x_1, x_2, x_0), the equations of the absolutes of elliptic geometry, hyperbolic geometry (and the other two geometries with the same absolute), Euclidean geometry, pseudo-Euclidean geometry, anti-Euclidean geometry, antipseudo-Euclidean geometry, and Galilean geometry can be written as:

$$x_1^2 + x_2^2 + x_0^2 = 0;\ x_1^2 + x_2^2 - x_0^2 = 0;\ x_0^2 = 0,\ x_1^2 + x_2^2 = 0;$$

$$x_0^2 = 0,\ x_1^2 - x_2^2 = 0;\ x_1^2 + x_2^2 = 0;\ x_1^2 - x_2^2 = 0;\ x_0^2 = 0,\ x_1^2 = 0;$$

however, a complete description of the absolute needs an indication of the lines belonging to it [a description of the equation of the absolute in line (tangential) coordinates]. For this, see, for example, F. Klein, *Nichteuklidische Geometrie* (Chelsea, New York), 1960.

of this geometry we may take any real conic, but restrict ourselves to points inside the absolute and lines intersecting the absolute (Figure A3b).[8]

3. *Antihyperbolic geometry* $[-1, +1]$ (elliptic measure of distances, hyperbolic measure of angles). The absolute of this geometry coincides with the absolute of hyperbolic geometry, but the points are now all points of the projective plane lying outside the absolute and the lines are lines not intersecting the absolute[9] (Figure A3c).[8]

4. *Doubly hyperbolic geometry* $[+1, +1]$ (hyperbolic measure of distances and angles). The absolute of this geometry coincides with the absolute of hyperbolic geometry, but the points are the points of the projective plane lying outside the absolute and the lines are lines intersecting the absolute[9] (Figure A3d).[8]

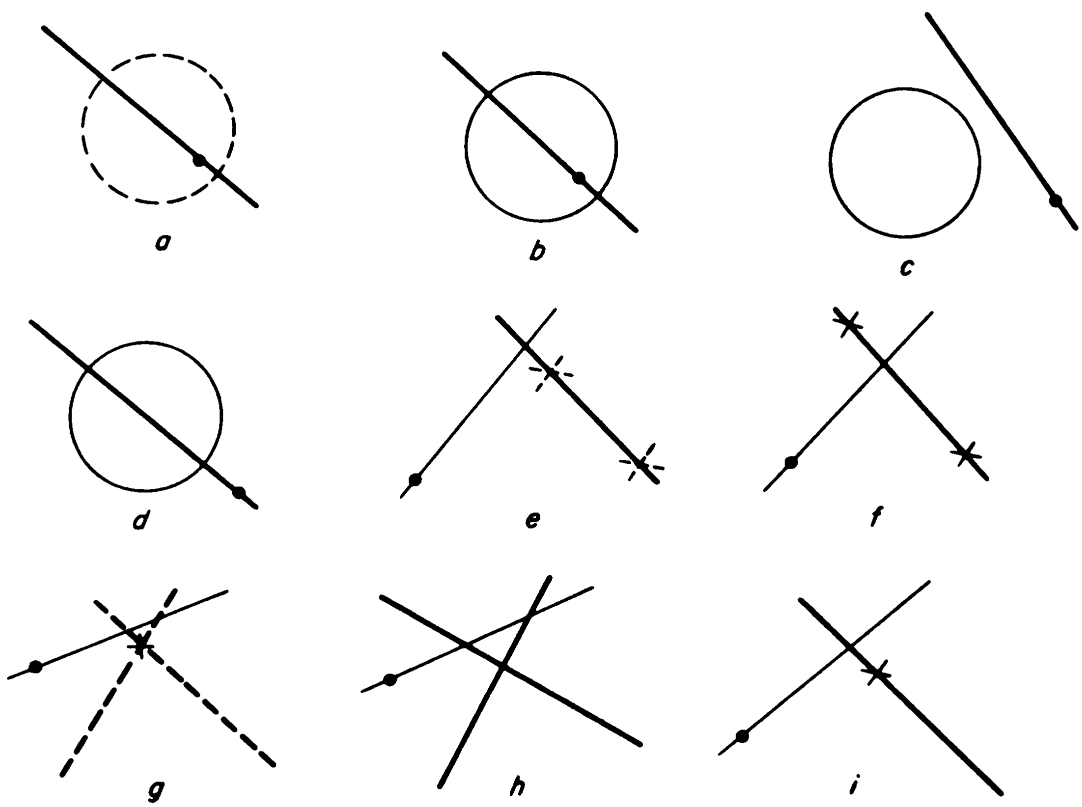


Fig. A3

[9] Thus is this geometry it may not be possible to join two points by a line (just as in the non-Euclidean geometry of Lobachevskii, for example, two lines may not have a point of intersection).

5. *Euclidean geometry* $[0, -1]$ (parabolic measure of distances, elliptic measure of angles). For the absolute of this geometry we may take any line of the (projective) plane (whose points are excluded henceforth) and a pair of conjugate complex points on this line (Figure A3e).[8]

6. *Pseudo-Euclidean geometry of Minkowski*[10] $[0, +1]$ (parabolic measure of distances, hyperbolic measure of angles). For the absolute of this geometry we may take any line (the line at infinity of the pseudo-Euclidean plane) and a pair of points of this line (Figure A3f); by points of pseudo-Euclidean geometry we understand all points of the projective plane not lying on the line at infinity, and by lines we understand all lines which intersect the line at infinity in points of some fixed segment bounded by the absolute points.[8]

7. *Anti-Euclidean geometry* $[-1, 0]$ (elliptic measure of distances, parabolic measure of angles). The absolute of this geometry consists of two conjugate complex lines intersecting at a (real) point of the projective plane (Figure A3g); this point and all lines through it are excluded.[8]

8. *Antipseudo-Euclidean geometry* $[+1, 0]$ (hyperbolic measure of distances, parabolic measure of angles). The absolute of this geometry consists of two (intersecting) lines; the points are all points of the projective plane belonging to a certain pair of vertically opposite angles formed by the absolute lines, and the lines are all lines not passing through the point of intersection of the absolute lines (Figure A3h).[8]

9. *Galilean geometry*[11] $[0, 0]$ (parabolic measure of distances and angles). The absolute of this geometry consists of a line and a point on it; the points are all points not lying on the absolute

---

[10] The famous German mathematician and physicist H. Minkowski (1864–1909) used this geometrical system for the geometrical description of phenomena which occur in the (special) theory of relativity.

[11] This name is explained by the close connection between the motions of this geometry and the so-called Galilean transformations in classical kinematics.

line and the lines are all lines not passing through the absolute point (Figure A3i).[8]

If the measure of distances is elliptic or hyperbolic, the distance between points A and B can be defined in all cases by the one formula A1, where I and J are the (real or conjugate complex) points of intersection of the line AB with the absolute of the geometry; if the measure of distances is parabolic, the line AB intersects the absolute in a unique point $I \equiv J$, and on this line we can introduce a Euclidean metric. If the measure of angles is elliptic or hyperbolic, the angle between two lines a and b can be defined in all cases by a common formula A1a, where i and j are the (real or conjugate complex) lines of the absolute which pass through the point of intersection of a and b (for example, the lines joining the vertex O of the angle to the absolute points in the cases of Euclidean and pseudo-Euclidean geometries); if the measurement of angles is parabolic, then through the point of intersection O of the lines a and b there passes a unique line belonging to the absolute, which enables us to define a parabolic (Euclidean) metric in the pencil of lines with vertex O. *Motions* are defined in all cases as point transformations which preserve distances and, also, take lines into lines and preserve angles between lines. If we start from the given description of all nine plane non-Euclidean geometries in the projective plane, the motions of these geometries will be represented by the projective transformations of the plane which preserve the absolute.

We note that the hyperbolic and antihyperbolic geometries, the Euclidean and anti-Euclidean geometries, the pseudo-Euclidean and antipseudo-Euclidean geometries, can be obtained from each other by simply interchanging points and lines: by agreeing to call the lines of Euclidean geometry "points" and the points "lines," angles between lines of the Euclidean plane "distances between points" and distances between points "angles between lines," we arrive at anti-Euclidean geometry. This fact is often denoted by saying that hyperbolic and anti-hyperbolic geometries, Euclidean and anti-Euclidean geometries, pseudo-Euclidean and antipseudo-Euclidean geometries, are *dual*

to each other. The other three plane non-Euclidean geometries, elliptic, doubly hyperbolic, and Galilean, are *self-dual*; this means that by calling the lines of, say, elliptic geometry "points" and vice versa, angles between lines "distances between points" and vice versa, we again arrive at elliptic geometry.

A2. Complex Coordinates of Points and Lines of the Plane Non-Euclidean Geometries

The results of Sections 7 and 9 of Chapter II can be thought of as introducing into the *ordinary* (*Euclidean*) plane certain complex coordinates of points and complex coordinates of (oriented) lines, connected with the polar coordinates r, φ of points and the polar coordinates θ, s of lines (cf. Figures A4a, b with Figures 1 and 28 on pp. 27 and 81) by the formulae

$$z = r(\cos\varphi + i\sin\varphi), \qquad z = \tan\frac{\theta}{2}(1 + \varepsilon s) \qquad \text{(A2a,b)}$$

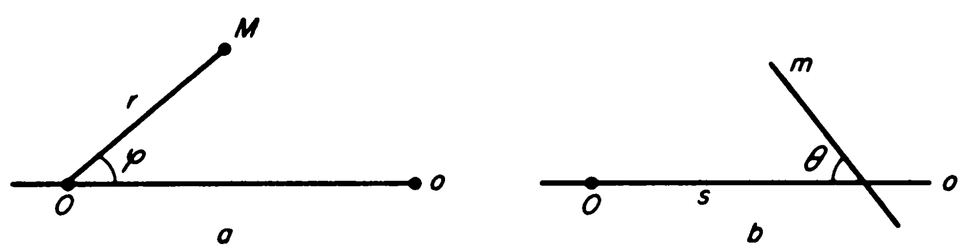


Fig. A4

By virtue of the duality of Euclidean and anti-Euclidean geometries, the same results can be interpreted as introducing complex coordinates of (oriented) points and lines into *anti-Euclidean* geometry, connected with the polar coordinates r, φ and θ, s of points and lines (see Figure A4) by the formulae

$$z = \tan\frac{r}{2}(1 + \varepsilon\varphi), \qquad z = \theta(\cos s + i\sin s) \qquad \text{(A2c,d)}$$

Similarly the results of Sections 11 and 12 of Chapter II show how it is possible to introduce complex coordinates of (oriented) points and (oriented) lines into the hyperbolic plane of Lobachevskii and the antihyperbolic plane dual to it; these

coordinates are connected with the polar coordinates r, φ of points and the polar coordinates θ, s of lines by the formulae

$$z = \tanh\frac{r}{2}(\cos\varphi + i\sin\varphi), \qquad z = \tan\frac{\theta}{2}(\cosh s + e\sinh s) \tag{A2e,f}$$

in the case of *hyperbolic* geometry, and by the formulae

$$z = \tan\frac{r}{2}(\cosh\varphi + e\sinh\varphi), \qquad z = \tanh\frac{\theta}{2}(\cos s + i\sin s) \tag{A2g,h}$$

in the case of *antihyperbolic* geometry. Finally, the mapping of the set of points of the *elliptic* plane of Riemann onto the set of complex numbers is very well known.[12] This mapping can be described geometrically as follows: We map the set of (oriented) points of the elliptic plane onto the set of points of an ordinary sphere σ of diameter 1 with equation $x^2 + y^2 + z^2 - z = 0$; two diametrically opposite points of the sphere are mapped onto the same point of the elliptic plane, taken *with opposite orientations*. Further, we map the sphere stereographically onto a plane π, which we take as the plane of the complex variable $z = x + iy$ (Figure A5). Analytically this mapping is given by the formula

$$z = \tan\frac{r}{2}(\cos\varphi + i\sin\varphi) \tag{A2i}$$

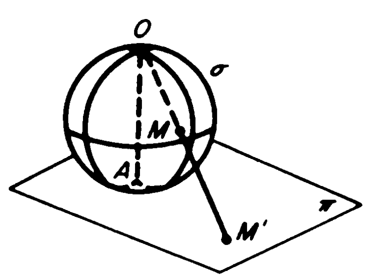


FIG. A5

where r, φ are the polar coordinates of a point of the elliptic plane (Figure A4a); oppositely oriented points correspond to complex numbers z and z_1 such that $z\bar{z}_1 = -1$ (that is, the equation $z' = -1/\bar{z}$ describes a *reorientation* of the elliptic plane; see

[12] See, for example Section 15 of Schwerdtfeger's book referred to on p. 109.

p. 111). The duality between points and lines of the elliptic plane enables us to interpret these results as a mapping

$$z = \tan\frac{\theta}{2}(\cos s + i \sin s) \tag{A2j}$$

of the set of (oriented) lines of the elliptic plane onto the set of complex numbers; here θ, s are the polar coordinates of (ordinary) lines of the elliptic plane (Figure A4b).

A mapping of the set of points and (oriented) lines of the *pseudo-Euclidean* plane of Minkowski onto the set of complex (more precisely, double and dual) numbers is effected by means of the formulae

$$z = r(\cosh\varphi + e \sinh\varphi), \qquad z = \tanh\frac{\theta}{2}(1 + \varepsilon s) \tag{A2k,l}$$

very close to formulae A2a,b. The points of the isotropic lines (lines at infinity) of the pseudo-Euclidean plane passing through the pole O of the system of polar coordinates (lines joining O to the absolute points; any segment of such lines has zero length)[13] correspond to divisors of zero, that is, to double numbers of the form $z = a(1 \pm e)$. In certain cases, for example, for a study of circular transformations of the pseudo-Euclidean plane, with which we shall deal below, it proves advisable to add "ideal points," or points at infinity, related to the point at infinity ∞ of the Euclidean plane; these (many) points at infinity correspond to the "ideal" double numbers $c\omega_1$, $c\omega_2$, $\sigma_1 = (1-e)/(1+e)$, $\sigma_2 = (1+e)/(1-e)$, ∞. Under a mapping of the (oriented) lines of the pseudo-Euclidean plane onto the set of dual numbers by formula A2l, just as in Section 9 we must associate lines parallel to the polar axis o, and oriented in the same way, with dual numbers $b\varepsilon$ of zero modulus; lines parallel to the polar axis o and oriented in the opposite way to it are associated with the

[13] The pseudo-Euclidean plane can be described as a plane (x, y) (where x, y are ordinary Cartesian coordinates) with metric

$$d_{AB}^2 = (x_2 - x_1)^2 - (y_2 - y_1)^2$$

[here (x_1, y_1) and (x_2, y_2) are the coordinates of the points A and B]; the isotropic lines through the pole O have equations $x - y = 0$ and $x + y = 0$.

ideal dual numbers $c\omega$. To the isotropic lines, which are naturally introduced as lines at infinity of the pseudo-Euclidean plane (these lines make an infinitely large angle with any line of the pseudo-Euclidean plane), correspond dual numbers of unit modulus (numbers z such that $z\bar{z} = 1$); these lines have no orientation. It is clear that formulae A2k,l can also be interpreted as mappings of points and lines of the *antipseudo-Euclidean* plane onto the set of dual numbers and double numbers by means of the formulae

$$z = r(1 + \varepsilon\varphi), \qquad z = \theta(\cosh s + e \sinh s) \qquad \text{(A2m,n)}$$

We now agree to take as absolute of *doubly hyperbolic* geometry a unit circle S of the Euclidean plane; as polar axis we take the horizontal diameter o of this circle, and as pole the point at infinity O of the polar axis (Figure A6). The polar coordinates r, φ of a point M are defined by the formulae

$$\begin{aligned} r &= \log W(M, O; I, J) = \log \frac{MI}{MJ}, \\ \varphi &= \log W(MO, o; i, j) = \log \frac{M_0 I_0}{M_0 J_0} \end{aligned} \qquad \text{(A3a,b)}$$

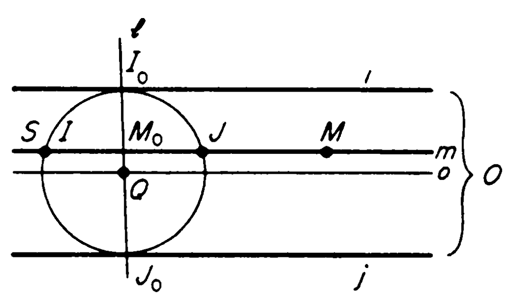


FIG. A6

where $W(M, O; I, J)$ and $W(MO, o; i, j)$ are the cross-ratios of four points and four lines respectively, I and J are the points of intersection of the line MO with the absolute (these points may be conjugate complex), i and j are the tangents of the absolute parallel to o (tangents drawn from the point O), and Q, M_0, I_0, J_0 are the points of intersection of the diameter l of S perpendicular to o with the lines o, MO, i, j. Hence it is easy to deduce that

under a mapping of the set of points of the doubly hyperbolic plane onto the set of double numbers, by means of the formula

$$z = \tanh\frac{r}{2}(\cosh\varphi + e\sinh\varphi) \qquad \text{(A2o)}$$

points of the doubly hyperbolic plane inside the band formed by the lines i and j correspond to double numbers

$$z = \tanh\frac{r}{2}(\cosh\varphi + e\sinh\varphi) = x + ey$$

for which $|y/x| = |\tanh\phi| < 1$, $x^2 - y^2 = \tanh^2 r/2 < 1$ (in the plane (x, y) of the double variable $z = x + ey$, these numbers correspond to points of the region I in Figure A7). As the point M

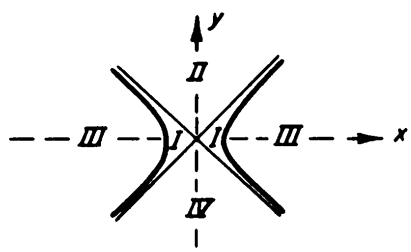


FIG. A7

approaches the lines i and j the magnitude of the coordinate φ increases without limit; hence, with points of the lines i and j we associate points of the asymptotes $y = \pm x$ of the hyperbola $x^2 - y^2 = 1$ in the plane of the double variable $z = x + ey$ [that is, to points of the lines i and j correspond divisors of zero $z = a(1 \pm e)$]. As the point M approaches the circle S, the value of the coordinate r increases without limit; hence it follows that with points of the absolute S (points at infinity of the doubly hyperbolic plane) we must associate points of the hyperbola $x^2 - y^2 = 1$ (points of the hyperbola $z\bar{z} = 1$) in the plane of the double variable $z = x + ey$. By the principle of continuity it is easy to establish that with the points I_0 and J_0 of Figure A6 we must associate the ideal double numbers σ_1 and σ_2. Finally, to points of the doubly hyperbolic plane lying outside the band formed by the lines i and j correspond values of the polar coordinate φ of the form $\varphi' + \frac{1}{2}\pi i$ and purely imaginary values $r = ir'$ of the coordinate r; hence it follows that with these

points are associated double numbers $x + ey$ for which $|y/x| > 1$ and $y > 0$ (region II of Figure A7). Finally, if we agree to regard all points of the doubly hyperbolic plane (except the points at infinity) as oriented and to regard two points differing only in orientation as corresponding to two double numbers z and z_1 such that $z\bar{z}_1 = 1$ (that is, we agree that a reorientation in the doubly hyperbolic plane is given by the equation $z' = 1/\bar{z}$) then it is easily seen that to points of the band formed by the lines i and j, oriented in the opposite way to the pole O, correspond double numbers $z = x + ey$ represented in Figure A7 by points of the region III, and to points outside this band, oriented in the opposite way to O, correspond numbers represented by points of the region IV; to points of the lines i and j, oriented in the opposite way to O, correspond ideal numbers $c\omega_1$ and $c\omega_2$, and to the antipole O', differing from O only in orientation, corresponds the number ∞. It is clear that by virtue of the duality between points and lines of the doubly hyperbolic plane all these results are easily interpreted as mappings of (oriented) lines of the doubly hyperbolic plane onto the set of double numbers by means of the formula

$$z = \tanh\frac{\theta}{2}(\cosh s + e\sinh s) \qquad \text{(A2p)}$$

where θ, s are the polar coordinates of a line (see Figure A4b).

Finally, we identify the set of points of the *Galilean* plane with points of an ordinary (Euclidean or affine) plane; that is, we regard the absolute line of this plane as the line at infinity, which, added to the affine plane, makes it projective; as absolute point we take the point at infinity of a line in the vertical direction, so that vertical lines are excluded. The formula

$$z = r(1 + \epsilon\varphi) \qquad \text{(A2q)}$$

defines a natural mapping of the points of the Galilean plane (Figure A8) onto the set of dual numbers; the set of lines of the Galilean plane is mapped onto the set of dual numbers by the analogous formula

$$z = \theta(1 + \epsilon s) \qquad \text{(A2r)}$$

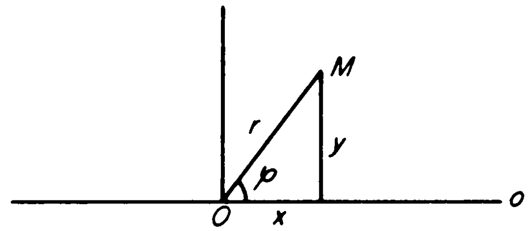


FIG. A8

In certain cases it is useful to adjoin to the set of points and lines of the Galilean plane ideal points and lines (points and lines at infinity), corresponding to the ideal dual numbers $c\omega$ and ∞ (the points at infinity of the Galilean plane can be interpreted as points of the absolute).

We now state, without proof, that *in the complex coordinates of points and lines described above, the distance* $d_{z_1z_2}$ *between two points* z_1 *and* z_2 *of the elliptic plane, the antihyperbolic plane, and the anti-Euclidean plane, and the angle* $\delta_{z_1z_2}$ *between two lines* z_1 *and* z_2 *of the elliptic plane, the hyperbolic plane, and the Euclidean plane are given by the same formulae*

$$\cos^2\frac{d}{2}=\frac{(1+z_1\bar{z}_2)(1+\bar{z}_1z_2)}{(1+z_1\bar{z}_1)(1+z_2\bar{z}_2)},\qquad \cos\frac{\delta}{2}=\frac{(1+z_1\bar{z}_2)(1+\bar{z}_1z_2)}{(1+z_1\bar{z}_1)(1+z_2\bar{z}_2)} \tag{A4a,b}$$

the distance $d_{z_1z_2}$ *between points* z_1 *and* z_2 *of the hyperbolic plane, the doubly hyperbolic plane, and the antipseudo-Euclidean plane, and the angle* $\delta_{z_1z_2}$ *between lines* z_1 *and* z_2 *of the antihyperbolic plane, the doubly hyperbolic plane, and the pseudo-Euclidean plane are given by the same formulae*

$$\cosh^2\frac{d}{2}=\frac{(1-z_1\bar{z}_2)(1-\bar{z}_1z_2)}{(1-z_1\bar{z}_1)(1-z_2\bar{z}_2)},\qquad \cosh\frac{\delta}{2}=\frac{(1-z_1\bar{z}_2)(1-\bar{z}_1z_2)}{(1-z_1\bar{z}_1)(1-z_2\bar{z}_2)} \tag{A4c,d}$$

finally, the distance $d_{z_1z_2}$ *between points* z_1 *and* z_2 *of the Euclidean plane, the pseudo-Euclidean plane, and the Galilean plane, and*

the angle $\delta_{z_1z_2}$ *between the lines* z_1 *and* z_2 *of the anti-Euclidean plane, the antipseudo-Euclidean plane, and the Galilean plane are given by the same formulae*

$$d^2 = (z_1 - z_2)(\bar{z}_1 - \bar{z}_2), \qquad \delta^2 = (z_1 - z_2)(\bar{z}_1 - \bar{z}_2) \quad \text{(A4e,f)}$$

In addition, *motions of the elliptic plane, the antihyperbolic plane, and the anti-Euclidean plane are given by the formulae*

$$z' = \frac{pz + q}{-\bar{q}z + \bar{p}}, \qquad p\bar{p} + q\bar{q} = 1 \quad \text{(A5a)}$$

motions of the hyperbolic plane, the doubly hyperbolic plane, and the antipseudo-Euclidean plane are given by the formulae

$$z' = \frac{pz + q}{\bar{q}z + \bar{p}}, \qquad p\bar{p} - q\bar{q} = 1 \quad \text{(A5b)}$$

finally, motions of the Euclidean plane, the pseudo-Euclidean plane, and the Galilean plane are given by the formulae

$$z' = pz + q, \qquad p\bar{p} = 1 \quad \text{(A5c)}$$

and by analogous formulae (cf. pp. 29 and 87–88). (It is clear that similar formulae in complex line coordinates give the motions in all nine plane non-Euclidean geometries.)

A3. Cycles and Circular Transformations

A *cycle* in all nine plane non-Euclidean geometries is defined as a curve (1-parameter set of points or lines) which admits a "sliding along itself" which allows us to connect any point of the cycle with any other point of the cycle. The cycles of *Euclidean* geometry are *circles* (in particular, *lines*), and the same *circles* regarded as sets of tangents (in particular, *points*), and also *pencils of parallel lines*. The cycles of the *non-Euclidean geometry of Lobachevskii* are *circles*, *horocycles*, and *equidistant curves* (in particular, *lines*), regarded as sets of points, and the same *circles* (in particular, *points*), *horocycles*, and *equidistant curves* regarded as sets of tangents, and, in addition, *pencils of equal inclination* (in particular, *orthogonal pencils*, or ideal points of the plane of

Lobachevskii), and *parallel pencils* (in particular, *points at infinity* of the Lobachevskii plane).[14] The cycles of *elliptic* geometry are *circles*, regarded as sets of points (in particular, *lines*, which can be regarded as circles of radius $\frac{1}{2}\pi$), and the same *circles*, regarded as sets of tangents (in particular, *points*, circles of radius 0).

In *doubly hyperbolic* geometry some of the cycles are *circles* (sets of points at a constant distance r, which can be real and positive, zero or purely imaginary, from a fixed point O); the point O is called the *center* of the circle, and the number r its *radius*. It is not difficult to see that a circle of radius 0 with center O consists of a pair of isotropic lines passing through O, and represented in the projective model of doubly hyperbolic geometry described above by tangents to the absolute S (Figure A9a);

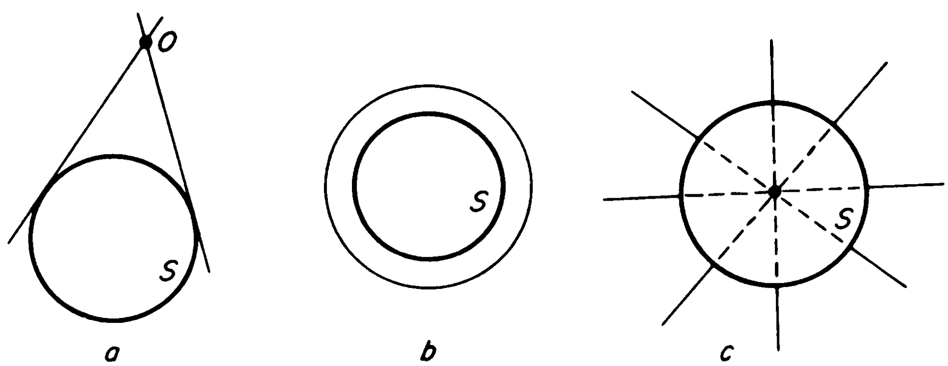


FIG. A9

in the limiting case, such a null circle is one *isotropic line* (counted twice) which can be called a line at infinity. A special case of a circle of imaginary radius is a *line*; thus, for example, it is easy to see that all points of the line l shown in Figure A6 are at the same (purely imaginary) distance from the point O. If the center of the circle which passes through the point M and touches the line m at this point is infinitely distant from M, then in the limit we obtain a *horocycle*, which also belongs to the cycles of doubly hyperbolic geometry. Finally, in doubly hyperbolic geometry there exists another type of cycle, characterized by the fact that a sliding along itself is a motion which does not leave

[14] See p. 124.

any point or *any* line of the plane invariant; an example of such a cycle, which can be called a *pseudocircle*, or ideal circle, is given in our model by a circle concentric with S (Figure A9b). A particular case of a pseudocircle is a *pseudoline*, or ideal line, represented in the projective model by a line *not intersecting* the absolute (for example, the line at infinity on Figure A9b). Since doubly hyperbolic geometry is self-dual, it is not difficult to obtain from this description a classification of cycles of the doubly hyperbolic plane regarded as sets of lines; among these cycles are the *points* of the doubly hyperbolic plane (that is, pencils of lines meeting in one point), *points at infinity* (points of the absolute or pencils of lines parallel in the sense of doubly hyperbolic geometry), and *ideal points*, represented on the projective model by points inside the absolute (Figure A9c); we note that every ideal circle of the doubly hyperbolic plane has an ideal center.

In *pseudo-Euclidean* geometry (in Figure A10 we regard the absolute of the geometry as consisting of the line at infinity of the affine plane and the two points of it whose directions are those of the bisectors of the angles between the coordinate axes) the cycles are *circles* of real (Figure A10a), zero (Figure A10c), or imaginary (Figure A10b) radius, and also the (ordinary,

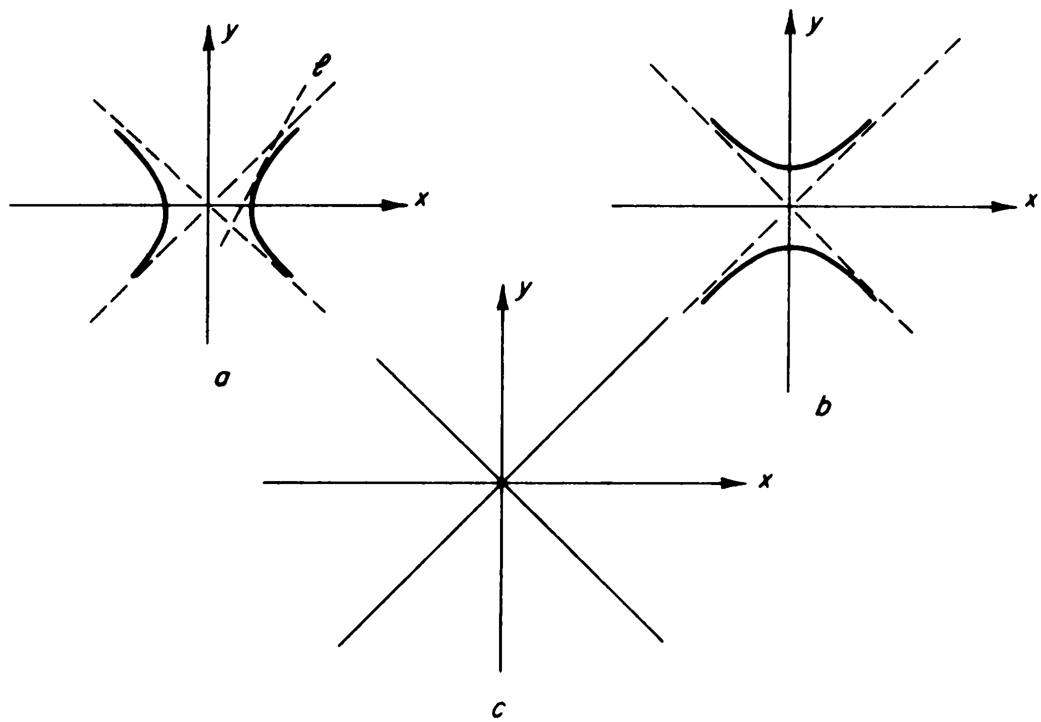


Fig. A10

isotropic, and ideal) *lines* of the pseudo-Euclidean geometry,[15] regarded as the sets of their points. Cycles, regarded as sets of lines of pseudo-Euclidean geometry, are *circles of imaginary radius*,[16] in particular, *points* (pencils of lines meeting in one point) and also *pencils of parallel lines.*

In *Galilean* geometry some of the cycles are *circles*, namely sets of points at a constant distance r (the *radius* of the circle) from a fixed point O (the *center* of the circle); in the projective (more precisely, affine) model of the Galilean plane considered in Figure A8 these circles are represented by pairs of vertical lines (Figure A11a); in the particular case $r = 0$ the circle reduces

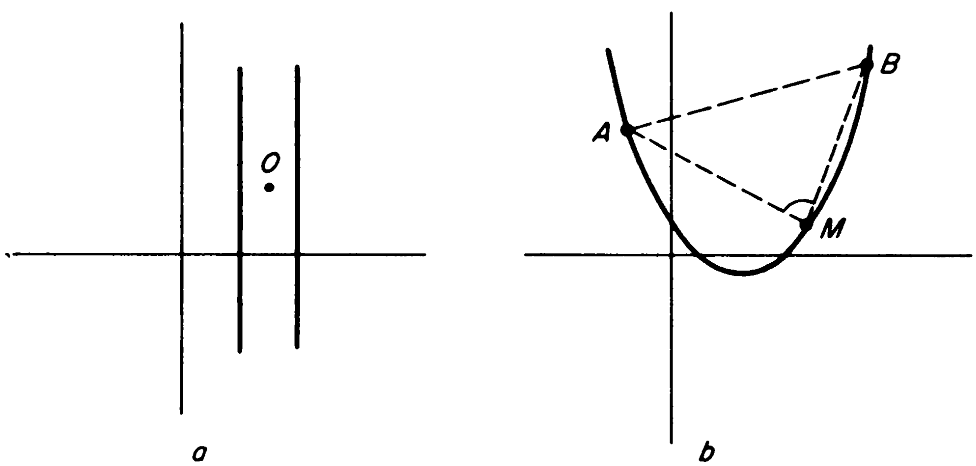


FIG. A11

to an *isotropic line* (or line at infinity) of the Galilean plane. Apart from circles, the is another type of cycle in the Galilean plane, which can be defined as the *set of points from which a given*

---

[15] If the pseudo-Euclidean plane is referred to the rectangular Cartesian coordinates x, y shown in Figure A10, the equations of the circles of this geometry have the form

$$a(x^2 - y^2) + 2b_1x + 2b_2y + c = 0$$

The particular case of a circle which we obtain by putting $a = 0$ is a line $2b_1x + 2b_2y + c = 0$ (which is regarded as a line of the pseudo-Euclidean plane only if $b_2^2 - b_1^2 > 0$).

[16] In the pseudo-Euclidean plane a circle of real radius does not have tangents (it is drawn by a dotted line in Figure A10a, and the "tangent" l is not a line of the pseudo-Euclidean geometry).

segment AB is seen at a constant (Galilean) *angle*; these cycles are represented in the same model by parabolas with vertical axes (Figure A11b). A particular case of the latter nondegenerate cycle of the Galilean plane is a *line*.[17] By virtue of the duality of the Galilean plane, we may obtain from this description the cycles of Galilean geometry regarded as sets of lines—these are *pairs of parallel pencils* (in particular, simply *parallel pencils*, or points at infinity) and *sets of tangents to the cycles represented in* Figure A11b (as a limiting case we have *pencils of intersecting lines*).

We shall not stop to describe the cycles of the antihyperbolic, anti-Euclidean, and antipseudo-Euclidean geometries, since these are dual to the hyperbolic, Euclidean, and pseudo-Euclidean geometries considered above.

It is remarkable that *for all nine plane non-Euclidean geometries the equation of any cycle, regarded as a set of points or a set of lines, can be written in terms of the complex point coordinates described above, or the complex line coordinates, by an equation of the form*[18]

$$Az\bar{z} + Bz - \bar{B}\bar{z} + C = 0, \quad A \text{ and } C \text{ purely imaginary} \quad \text{(A6)}$$

Equation A6 in the complex point coordinates *denotes a line in elliptic, antihyperbolic, and anti-Euclidean geometries if and only if*[19]

$$A + C = 0 \quad \text{(A7a)}$$

---

[17] In the coordinates x, y of the Galilean plane which occur in Figure A8, the equation of a nondegenerate cycle has the form

$$ax^2 + 2b_1x + 2b_2y + c = 0$$

If $a = 0$ we obtain the equation of a line, and if $C_2 = 0$, the equation of a circle.

[18] It is not difficult to establish the conditions imposed on the coefficients of Equation 6 which characterize the various types of cycles in each of the non-Euclidean geometries (cf. pp. 115 and 125); we shall leave the reader to do this.

[19] Here, among lines are included the isotropic lines (or lines at infinity) and ideal lines described above; we obtain an ordinary line if as well as condition (A7) the condition

$$AC + B\bar{B} > 0$$

is satisfied.

in hyperbolic, doubly hyperbolic, and antipseudo-Euclidean geometries if and only if [19]

$$A - C = 0 \tag{A7b}$$

in Euclidean, pseudo-Euclidean, and Galilean geometries if and only if [19]

$$A = 0 \tag{A7c}$$

[Condition (A7a) is also necessary and sufficient for Equation A6, written in complex *line coordinates* in the elliptic, hyperbolic, or Euclidean plane, to represent a *point*; the corresponding condition for the antihyperbolic, doubly hyperbolic, and pseudo-Euclidean planes has the form (A7b), and that for the anti-Euclidean, antipseudo-Euclidean, and Galilean planes has the form (A7c).]

The absolute is described in complex point coordinates in hyperbolic, doubly hyperbolic, and antipseudo-Euclidean geometries by the equation

$$z\bar{z} = 1 \tag{A8a}$$

and in antihyperbolic geometry by the equation

$$z\bar{z} = -1. \tag{A8b}$$

(The same Equation A8a describes the absolute in complex line coordinates in antihyperbolic, doubly hyperbolic, and pseudo-Euclidean geometries, and Equation A8b similarly in hyperbolic geometry.)

We may show that *in all nine plane non-Euclidean geometries circular transformations* [transformations in the set of points or lines (axes) which take each cycle into a cycle] *are described by the formulae*

$$\begin{gathered} z' = \frac{az + b}{cz + d} \quad \textit{or} \quad z' = \frac{a\bar{z} + b}{c\bar{z} + d}, \\ \Delta = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \quad \text{not a divisor of zero} \end{gathered} \tag{A9}$$

The circular point transformations are in all cases *conformal*, that is, they preserve angles between curves (see footnote 67 on

p. 137), and the axial circular transformations (transformations in the set of lines) are in all cases *equidistantial* (that is, they preserve tangential distances between curves; see footnote 80 on p. 162). This follows in particular from the following theorem: *All conformal transformations in any of the nine plane non-Euclidean geometries are described in complex point coordinates by the formula*

$$z' = f(z) \tag{A10}$$

where $f(z)$ is an analytic function (that is, it has a derivative) *of an ordinary complex, dual, or double variable*[20]; *the same formula* (A10) *for an analytic function $f(z)$ in complex line coordinates in the nine plane non-Euclidean geometries describes equidistantial transformations of the plane.* We shall not prove these statements.[21]

In conclusion we note that if we pass over from complex coordinates $z = x + iy$, $z = x + ey$, or $z = x + \varepsilon y$ to coor-

---

[20] That is, a function $f(z)$ such that the limit

$$\frac{df}{dz} = f'(z) = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}$$

has the same value irrespective of the path (in the plane of the complex variable of the corresponding type) along which Δz tends to zero. (It is understood that the prime in the number z' which occurs in Equation A10 *does not* denote the derivative.)

[21] An analytic function $z' = f(z)$ of a *dual* variable $z = x + \varepsilon y$ can be written as a function

$$f(x + \varepsilon y) = f(x) + \varepsilon y f'(x)$$

(since in the Taylor expansion of $f(z)$ in powers of εy, all the remaining terms vanish by virtue of the equation $\varepsilon^2 = 0$). Hence we may easily obtain the following explicit description of all equidistantial transformations of the Euclidean plane: *if a line $z = \tan \theta/2\,(1 + \varepsilon s)$ of the Euclidean plane is reduced to normal coordinates* (which occur in the normal equation of a line) *p and α*, where p is the (oriented) distance of a line from the pole O, and α is the angle which the perpendicular drawn from O to the line makes with the polar axis o, *then the transformation*

$$p' = F(p, \alpha), \quad \alpha' = f(p, \alpha)$$

is equidistantial if and only if

$$f(p, \alpha) = f_1(\alpha), \quad F(p, \alpha) = f_1'(\alpha)p + f_2(\alpha)$$

We shall leave the reader to obtain this result.

dinates x, y we may verify that in these coordinates Equation A6 of a cycle is written as

$$a(x^2 + y^2) + 2b_1x + 2b_2y + c = 0, \qquad \text{(A11a)}$$

$$a(x^2 - y^2) + 2b_1x + 2b_2y + c = 0 \qquad \text{(A11b)}$$

or

$$ax^2 + 2b_1x + 2b_2y + c = 0 \qquad \text{(A11c)}$$

respectively. Hence the cycles of elliptic, hyperbolic, and Euclidean geometries can be represented on the (x, y) plane as *circles*,[22] the cycles of antihyperbolic, doubly hyperbolic, and pseudo-Euclidean geometries as *hyperbolas* of the (x, y) plane, and finally the cycles of anti-Euclidean, antipseudo-Euclidean, and Galilean geometries by *parabolas* of the (x, y) plane; however, we must resort to a rather complicated completion of the plane by ideal (or infinite) elements, corresponding to the ideal complex numbers ∞; $c\omega$; $c\omega_1$, $c\omega_2$, σ_1, and σ_2. We immediately obtain the result that the set of all circular transformation of the elliptic, hyperbolic, and Euclidean planes can be represented as the set of all transformations of the (x, y) plane *which take the family* (Equation A11a) *of circles into itself*; the set of all circular transformations of the antihyperbolic, doubly hyperbolic, and pseudo-Euclidean planes as the set of all transformations of the (x, y) plane *which take the family* (Equation A11b) *of rectangular hyperbolas with fixed directions of axes into itself*; finally, the set of all circular transformations of the anti-Euclidean, antipseudo-Euclidean, and Galilean planes as the set of all transformations of the (x, y) plane *which take the family* (Equation A11c) *of parabolas with fixed direction of axis into itself.* Similar representations exist for axial circular transformations of the nine non-Euclidean planes.

---

[22] The corresponding representation of the elliptic and hyperbolic planes on the (x, y) plane is called the *Poincaré model* of these planes (see, for example, the appendix to Chapter II of Part III of I. M. Yaglom, *Geometric Transformations*, Random House, New York).

¶A, p. 22. In geometry and in algebra we also consider hypercomplex numbers of the form (44), whose coefficients a_0, a_1,..., a_n are themselves complex or hypercomplex numbers (for example, ordinary complex numbers, double numbers, or quaternions). We shall not go into these further generalizations of the idea of complex number.

¶B, p. 23. Some particular systems of hypercomplex numbers may be applied to geometry, for example, the so-called *alternions* (from the Latin *alternus*—alternating) or *Clifford numbers*, which have $2^n - 1$ complex units

$$e_i\,,\ e_ie_j = e_{ij}\,,\ e_ie_je_k = e_{ijk}\,,\ldots,\ e_1e_2 \ldots e_n = e_{12\ldots n} \qquad (i, j, k \ldots = 1, 2,\ldots, n)$$

with rules of multiplication for these units

$$e_i^2 = -1,\ e_ie_j = -e_je_i \quad \text{if} \quad i \neq j,$$

$$e_{i_1i_2\ldots i_k} \cdot e_{j_1j_2\ldots j_l} = e_{i_1i_2\ldots i_kj_1j_2\ldots j_l} = e_{i_1}e_{i_2} \ldots e_{i_k}e_{j_1}e_{j_2} \ldots e_{j_l}$$

where the product $e_{i_1}e_{i_2} \ldots e_{i_k}e_{j_1}e_{j_2} \ldots e_{j_l}$ can be transformed by means of the rules $e_i^2 = -1$, $e_ie_j = -e_je_i$ (by means of these rules each unit $e_{i_1}e_{i_2} \ldots e_{i_k} = e_{i_1i_2\ldots i_k}$ can be reduced to the form $e_{i'_1i'_2\ldots i'_{k'}}$, where $k' \leqslant k$ and $i'_1 < i'_2 < \cdots < i'_{k'}$; hence the number of different units is equal to $\binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n - 1$).* The definition of alternions can be somewhat generalized by putting $e_i^2 = \pm 1$, $i = 1, 2,\ldots, n$ (*generalized alternions*

* It is easily seen that if $n = 1$ alternions reduce to ordinary complex numbers, and if $n = 2$ they reduce to quaternions (see p. 24).

or *pseudoalternions*).† Recently, interesting applications have been found for the so-called ***plural numbers*** (from the Latin *plus*—many) with $n - 1$ complex units

$$\varepsilon_1 = \varepsilon, \quad \varepsilon_2 = \varepsilon^2, \ldots, \quad \varepsilon_{n-1} = \varepsilon^{n-1}$$

and rules of multiplication for these units

$$\varepsilon_i \varepsilon_j = \varepsilon^i \varepsilon^j = \varepsilon^{i+j}, \quad \varepsilon^n = 0;$$

for $n = 1$, plural numbers reduce to dual numbers.

¶c, p. 25. The multiplication tables for the complex units in these six systems are respectively

	$\dot{i}_1$	$\dot{i}_2$	$\dot{i}_3$	e_1	e_2	e_3	e_4
$\dot{i}_1$	-1	$\dot{i}_3$	$-\dot{i}_2$	e_2	$-e_1$	$-e_4$	e_3
$\dot{i}_2$	$-\dot{i}_3$	-1	$\dot{i}_1$	e_3	e_4	$-e_1$	$-e_2$
$\dot{i}_3$	$\dot{i}_2$	$-\dot{i}_1$	-1	e_4	$-e_3$	e_2	$-e_1$
e_1	$-e_2$	$-e_3$	$-e_4$	1	$-\dot{i}_1$	$-\dot{i}_2$	$-\dot{i}_3$
e_2	e_1	$-e_4$	e_3	$\dot{i}_1$	1	$\dot{i}_3$	$-\dot{i}_2$
e_3	e_4	e_1	$-e_2$	$\dot{i}_2$	$-\dot{i}_3$	1	$\dot{i}_1$
e_4	$-e_3$	e_2	e_1	$\dot{i}_3$	$\dot{i}_2$	$-\dot{i}_1$	1

	$\dot{i}_1$	$\dot{i}_2$	$\dot{i}_3$	ε_1	ε_2	ε_3	ε_4
$\dot{i}_1$	-1	$\dot{i}_3$	$-\dot{i}_2$	ε_2	$-\varepsilon_1$	$-\varepsilon_4$	ε_3
$\dot{i}_2$	$-\dot{i}_3$	-1	$\dot{i}_1$	ε_3	ε_4	$-\varepsilon_1$	$-\varepsilon_2$
$\dot{i}_3$	$\dot{i}_2$	$-\dot{i}_1$	-1	ε_4	$-\varepsilon_3$	ε_2	$-\varepsilon_1$
ε_1	$-\varepsilon_2$	$-\varepsilon_3$	$-\varepsilon_4$	0	0	0	0
ε_2	ε_1	$-\varepsilon_4$	ε_3	0	0	0	0
ε_3	ε_4	ε_1	$-\varepsilon_2$	0	0	0	0
ε_4	$-\varepsilon_3$	ε_2	ε_1	0	0	0	0

† The number of numbers e_i^2 equal to $+1$ is called the *index* of the algebra of pseudoalternions. If $n = 1$ the index can be equal only to 0 or 1, and if $n = 2$ it can be equal to 0, 1, or 2; if $n = 1$, alternions of index 0 are ordinary complex numbers, and pseudoalternions of index 1 are double numbers; if $n = 2$, alternions of index 0 are quaternions and pseudoalternions of index 1 or 2 are pseudoquaternions (see p. 24; if $e_1^2 = -1$, $e_2^2 = +1$, then $(e_{12})^2 = (e_1e_2)^2 = e_1e_2e_1e_2 = -e_1^2e_2^2 = +1$, and if $e_1^2 = e_2^2 = +1$, then $(e_{12})^2 = (e_1e_2)^2 = e_1e_2e_1e_2 = -e_1^2e_2^2 = -1$).

	i	e_1	e_2	ε_1	ε_2	ε_3	ε_4
i	-1	e_2	$-e_1$	ε_2	$-\varepsilon_1$	$-\varepsilon_4$	ε_3
e_1	$-e_2$	1	$-i$	ε_3	ε_4	ε_1	ε_2
e_2	e_1	i	1	ε_4	$-\varepsilon_3$	$-\varepsilon_2$	ε_1
ε_1	$-\varepsilon_2$	$-\varepsilon_3$	$-\varepsilon_4$	0	0	0	0
ε_2	ε_1	$-\varepsilon_4$	ε_3	0	0	0	0
ε_3	ε_4	$-\varepsilon_1$	ε_2	0	0	0	0
ε_4	$-\varepsilon_3$	$-\varepsilon_2$	$-\varepsilon_1$	0	0	0	0

	i	ε_1	ε_2	ε_3	ε_4	ε_5	ε_6
i	-1	ε_2	$-\varepsilon_1$	ε_4	$-\varepsilon_3$	$-\varepsilon_6$	ε_5
ε_1	$-\varepsilon_2$	0	0	ε_5	ε_6	0	0
ε_2	ε_1	0	0	ε_6	$-\varepsilon_5$	0	0
ε_3	$-\varepsilon_4$	$-\varepsilon_5$	$-\varepsilon_6$	0	0	0	0
ε_4	ε_3	$-\varepsilon_6$	ε_5	0	0	0	0
ε_5	ε_6	0	0	0	0	0	0
ε_6	$-\varepsilon_5$	0	0	0	0	0	0

	e	ε_1	ε_2	ε_3	ε_4	ε_5	ε_6
e	1	ε_2	ε_1	ε_4	ε_3	$-\varepsilon_6$	$-\varepsilon_5$
ε_1	$-\varepsilon_2$	0	0	ε_5	ε_6	0	0
ε_2	$-\varepsilon_1$	0	0	ε_6	ε_5	0	0
ε_3	$-\varepsilon_4$	$-\varepsilon_5$	$-\varepsilon_6$	0	0	0	0
ε_4	$-\varepsilon_3$	$-\varepsilon_6$	$-\varepsilon_5$	0	0	0	0
ε_5	ε_6	0	0	0	0	0	0
ε_6	ε_5	0	0	0	0	0	0

	ε_1	ε_2	ε_3	ε_4	ε_5	ε_6	ε_7
ε_1	0	ε_3	0	ε_5	0	$-\varepsilon_7$	0
ε_2	$-\varepsilon_3$	0	0	ε_6	ε_7	0	0
ε_3	0	0	0	ε_7	0	0	0
ε_4	$-\varepsilon_5$	$-\varepsilon_6$	$-\varepsilon_7$	0	0	0	0
ε_5	0	$-\varepsilon_7$	0	0	0	0	0
ε_6	ε_7	0	0	0	0	0	0
ε_7	0	0	0	0	0	0	0

and the definition of conjugate is

$$\bar{Z} = a_0 - a_1E_1 - a_2E_2 - a_3E_3 - a_4E_4 - a_5E_5 - a_6E_6 - a_7E_7$$

if

$$Z = a_0 + a_1E_1 + a_2E_2 + a_3E_3 + a_4E_4 + a_5E_5 + a_6E_6 + a_7E_7$$

where the $E_1, E_2, E_3, E_4, E_5, E_6, E_7$ are the complex units which appear in the multiplication tables (taken in the same order as in the tables), and the definition of modulus is

$$|Z|^2 = Z\bar{Z} = \begin{cases} a_0^2 + a_1^2 + a_2^2 + a_3^2 - a_4^2 - a_5^2 - a_6^2 - a_7^2 & \text{for pseudooctaves} \\ a_0^2 + a_1^2 + a_2^2 + a_3^2 & \text{for degenerate octaves} \\ a_0^2 + a_1^2 - a_2^2 - a_3^2 & \text{for degenerate pseudooctaves} \\ a_0^2 + a_1^2 & \text{for doubly degenerate octaves} \\ a_0^2 - a_1^2 & \text{for doubly degenerate pseudooctaves} \\ a_0^2 & \text{for triply degenerate octaves} \end{cases}$$

¶D, p. 141. This theorem is sometimes called the ***basic theorem*** *of the theory of circular transformations.* From the basic theorem it follows that all circular transformations of the plane other than similarities reduce, in a certain sense, to a (unit) inversion*; this determines the important role which such a special transformation as inversion plays in the general theory of circular transformations. It may also be shown that *every circular transformation of the plane may be represented as the product of a finite number* (not more than six) *of symmetries with respect to circles* (*inversions*) *and symmetries about lines.*† In fact any similarity may be expressed as the product of a dilatation (whose center can be chosen arbitrarily, and whose ratio is

* See footnote 71 on p. 143.

† Since, in questions connected with circular transformations, a line is usually considered as a particular case of a circle ("circle of infinite radius"), a symmetry about a line can be regarded as a particular case of a symmetry with respect to a circle (inversion).

positive) and a motion[‡]; on the other hand, any motion can be expressed as the product of not more than three symmetries about lines,[§] and the dilatation $z' = kz$ with center O and ratio $k > 0$ can naturally be split up into the product of two inversions $z_1 = 1/\bar{z}$ and $z' = k/\bar{z}_1$. Hence every circular transformation—the product of a similarity and an inversion—can be expressed as the product of not more than $3 + 2 + 1 = 6$ inversions.

¶E, p. 141. The inversion (Equation 16) of power k takes the point z into a point z' such that the rays Oz and Oz' lie along the same line and $\{O, z\} \cdot \{O, z'\} = k$ (here $\{O, z\}$ and $\{O, z'\}$ are *oriented* distances; the orientation of the line Oz is chosen arbitrarily). The symmetry (Equation 16) with respect to the circle $z\bar{z} = k$ (where $k > 0$) can be defined by the property that every circle passing through two corresponding points z and z' is orthogonal to the circle $z\bar{z} = k$ (see Figure 56a).

¶F, p. 145. This is like the situation in ordinary (school, or Euclidean) geometry, where there is no advantage in shifting a figure used in the proof of a theorem to a new position, since the new figure does not differ from the original (it has the same geometrical properties).*

¶G, p. 147. In other words, in circular geometry (see the end of Section 13) every (convex) quadrangle is equal to some parallelogram; a quadrangle inscribed in a circle is equal to a rectangle, and a quadrangle in which the products of opposite sides are equal is equal to a rhombus. Hence the set of all (convex) quadrangles $\overline{z_1z_2z_3z_4}$ which are different from the point of view

---

‡ See, for example, I. M. Yaglom, *Geometric Transformations*, Part II (Random House, New York), 1966, and H. S. M. Coxeter, *Introduction to Geometry*, Chap. V (John Wiley, New York), 1961.

§ See, for example, I. M. Yaglom, *Geometric Transformations* (Random House, New York), 1962.

* In school, in proving theorems about, say, triangles, we always tend to draw a triangle so that one of its sides is horizontal; however, such a figure is only visually, but not mathematically, simpler than that obtained from it by a rotation.

of circular geometry (where we understand by a quadrangle simply a tetrad of points z_1, z_2, z_3, z_4 of the plane of the complex variable) can be interpreted as the set of all possible parallelograms not similar to each other (for example, the set of all parallelograms, one of whose sides is equal to 1, and another greater than or equal to 1; it is not difficult to see that two parallelograms not similar to each other are not equal to each other from the point of view of circular geometry).

¶H, p. 157. If $k < 0$, the circles (and lines) given by Equation 19a exhaust all the "circles" (in the sense of circular geometry, that is, circles and lines) which are taken into themselves by the inversion (Equation 16). However, if $k > 0$, then as well as these circles there exists one more circle $z\bar{z} = k$, which is also taken into itself by the inversion (Equation 16), and this circle does not have the form of Equation 19a; if it is written in the form of Equation 19, then $C/A = -k$. This circle plays a particular role with respect to the inversion: it is *pointwise invariant*, that is, the inversion takes each point of the circle into the *same* point, (and not into the second point of intersection of the line Oz with the circle, as in the case of a circle given by Equation 19a). The circle $z\bar{z} = k$ is sometimes called the **circle of inversion** of the transformation (Equation 16).

¶I p. 157. The inversion (Equation 16) takes an arbitrary circle (or line) (Equation 19) into a circle or line

$$Cz\bar{z} + Bkz - \bar{B}k\bar{z} + Ak^2 = 0 \tag{19b}$$

which is generally different from the original circle. If Equations 19 and 19b give two different circles S and S' (Figure 62A) then since each line l through O with equation $B_1 z - \bar{B}_1 \bar{z} = 0$ is taken by the inversion (Equation 16) into itself, the inversion takes the "circles" l and S into the "circles" l and S'. And since every circular transformation preserves the angles between circles, *each line l through O makes the same angles with the circles S and S'*. [More precisely, if l intersects the circle S at points z_1 and z_2 and the inversion (Equation 16) takes these

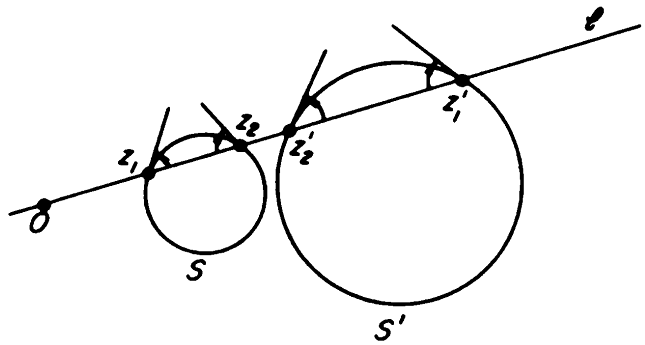


FIG. 62A

points into points z_1' and z_2' of the circle S', then the (oriented) angles $\{lz_1S\}$ and $\{lz_1'S'\}$, $\{lz_2S\}$ and $\{lz_2'S'\}$ differ in sign, and so the (oriented) angles $\{lz_1S\}$ and $\{lz_2'S'\}$, $\{lz_2S\}$ and $\{lz_1'S'\}$ are equal to each other.] Thus *the center O of the inversion which takes the circle S into the circle S' is a* ***center of similitude*** *of the circles S and S'** (see p. 104).

This property has interesting applications; e.g., let S be an arbitrary circle, and $z_1, z_2, z_3, z_4, z_5, z_6$ six arbitrary points of this circle, the vertices of a hexagon $\overline{z_1z_2z_3z_4z_5z_6}$ inscribed in the circle (this can be self-intersecting: see Figure 62B). We denote by Q_1, Q_2, and Q_3 the points of intersection of the opposite sides $[z_1z_2]$ and $[z_4z_5]$, $[z_2z_3]$ and $[z_5z_6]$, $[z_3z_4]$ and $[z_6z_1]$ of the hexagon $\overline{z_1z_2z_3z_4z_5z_6}$.† Since $\{Q_1z_1\}\cdot\{Q_1z_2\}=\{Q_1z_4\}\cdot\{Q_1z_5\}=k_1$

---

* If the inversion with center O and power k takes the points z_1 and z_2 of a circle S into the points z_1' and z_2' of a circle S' and the power of O with respect to S is equal to k_1, then obviously

$$|z_1'| = k/|z_1|, \quad |z_2'| = k/|z_2|, \quad \text{and} \quad |z_1|\cdot|z_2| = k_1$$

so with no difficulty we may deduce that

$$|z_1'| = \frac{k}{k_1}|z_2|, \quad |z_2'| = \frac{k}{k_1}|z_1|, \quad \text{and} \quad z_1' = \frac{k}{k_1}z_2, \quad z_2' = \frac{k}{k_1}z_1$$

thus *the dilatation with center O and ratio k/k_1 takes the circle S into the circle S'* (it takes the points z_1 and z_2 of S into the points z_2' and z_1' of S').

† For simplicity we assume here that no two opposite sides of the hexagon $\overline{z_1z_2z_3z_4z_5z_6}$ inscribed in the circle S are parallel; we leave the reader to explain how to modify the theorem proved here (Pascal's theorem) if this condition is not satisfied.

(the power of Q_1 with respect to S), the inversion with center Q_1 and power k_1 takes the points z_1 and z_4 into the points z_2 and z_5, and vice versa; hence it takes S into itself, and an arbitrary circle S_1, intersecting S at z_1 and z_4, into a circle S_2, intersecting S at z_2 and z_5. Further, $\{Q_2z_2\} \cdot \{Q_2z_3\} = \{Q_2z_5\} \cdot \{Q_2z_6\} = k_2$ (the

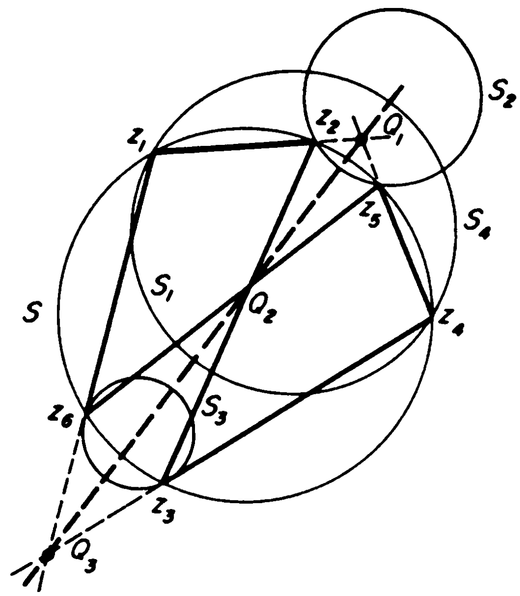


FIG. 62B

power of Q_2 with respect to S); hence the inversion with center Q_2 and power k_2 takes the points z_2 and z_5 into the points z_3 and z_6, the circle S into itself, and the circle S_2 into a circle S_3 intersecting S at z_3 and z_6. Finally, $\{Q_3z_3\} \cdot \{Q_3z_4\} = \{Q_3z_6\} \cdot \{Q_3z_1\} = k_3$ (the power of Q_3 with respect to S); hence the inversion with center Q_3 and power k_3 takes the circle S into itself, and the circle S_3 into a circle S_4, which intersects S at the same points z_4 and z_1 as S_1 does. Since each inversion changes the sense of the angle between circles, and the (oriented) angles between circles, measured at their two points of intersection, have opposite signs (see footnote 66 on p. 137), then

$$\{S_1z_1S\} = -\{S_2z_2S\} = \{S_3z_3S\} = -\{S_4z_4S\} = \{S_4z_1S\}$$

But from the equation

$$\{S_1z_1S\} = \{S_4z_1S\}$$

it follows that the circles S_1 and S_4 coincide (since there exists *only one* circle S_1 intersecting the given circle S at the fixed points z_1 and z_4 and making a known angle $\{S_1 z_1 S\} = \alpha$ with S at these points).

Thus we conclude that the points Q_1, Q_2, and Q_3 are the centers of the inversions which take S_1 into S_2, S_2 into S_3, and S_3 into S_1; that is, Q_1, Q_2, and Q_3 are the *centers of similitude* of the circle S_1 and S_2, S_2 and S_3, S_3 and S_1. But we know that the three centers of similitude of pairs of three circles lie on one line, the axis of similitude of the three circles (see p. 105). Hence it follows that *if* $\overline{z_1 z_2 z_3 z_4 z_5 z_6}$ *is an arbitrary hexagon* (possibly self-intersecting) *inscribed in a circle, the points of intersection of opposite sides lie on one line.* This theorem (more precisely, a somewhat more general theorem) was first proved in 1640 by the famous French mathematician, physicist, author, and philosopher B. Pascal (1623–1662) at the age of 16; it is therefore called *Pascal's theorem.*

We return to the idea of a bundle of circles (and lines)

$$Az\bar{z} + Bz - \bar{B}\bar{z} + Ak = 0 \tag{19a}$$

with center O and power k. Since the value of the ratio C/A of the coefficients of Equation 19 of a circle is equal to the power of the circle (the power of a point with respect to the circle), the bundle (Equation 19a) can be defined as *the set of all circles whose power is equal to the power k of the bundle* (that is, circles such that the power of the center O of the bundle with respect to them is equal to k) and all lines through the center O. [The bundle with center O can also be defined as the set of all circles such that the *radical center of any three of them* (see p. 48) *coincides with* fixed point O (and all lines through O).] The bundle of circles with arbitrary center w and power k can be defined as the set of all circles (Equation 19) such that

$$Aw\bar{w} + Bw - \bar{B}\bar{w} + (C - Ak) = 0 \tag{19'a}$$

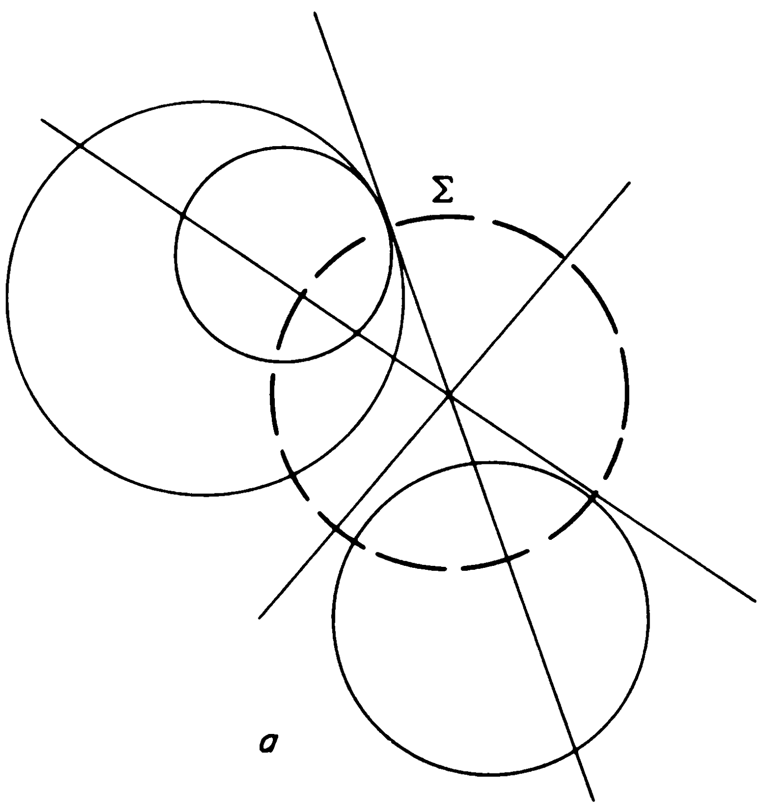
(see Equation 20, p. 46). Hence we may deduce purely analytically that *every circular transformation* (Equation 1 or 1a) *takes a bundle of circles into a bundle of circles* (that is, the idea of a bundle of circles *belongs to circular geometry*); however, the presence in the set of complex numbers

of the singular number ∞ leads to the necessity of somewhat extending the idea of "bundle of circles" by introducing the so-called "singular bundles," which we shall describe below.

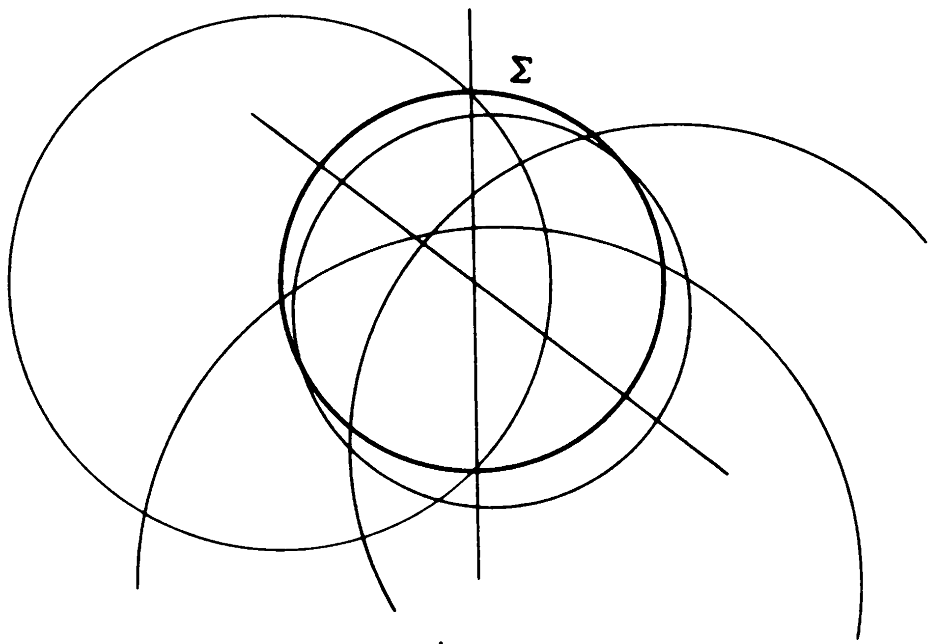
If the power k of the bundle (Equation 19a, or Equation 19 with condition 19′a) is positive, then the bundle is called *hyperbolic.* It is clear that the length Oz_0 of the segment of the tangent drawn from O to an arbitrary circle of the bundle is equal to $\sqrt{k}$ (see p. 44); hence it follows that a *hyperbolic bundle is formed by all circles* (and lines) *orthogonal to a fixed circle* Σ with center at O and radius $\sqrt{k}$ (the circle $z\bar{z} = k$; Figure 62C,a). Since circular transformations preserve angles between circles, it is clear that circular transformations take a hyperbolic bundle into a hyperbolic bundle; however, in order that this statement should be precise, it is necessary to include among hyperbolic bundles sets of all circles and lines orthogonal to a fixed line l (sets of all lines perpendicular to any fixed line l and all circles whose centers lie on l), since a circular transformation can take a circle Σ into a line l. The set of all circles (and lines) orthogonal to l is called a *singular hyperbolic bundle* with axis l.

In Equation 19a (or 19′a) it is possible that $k = 0$; in this case we go over to the set of circles and lines (Equation 19), where

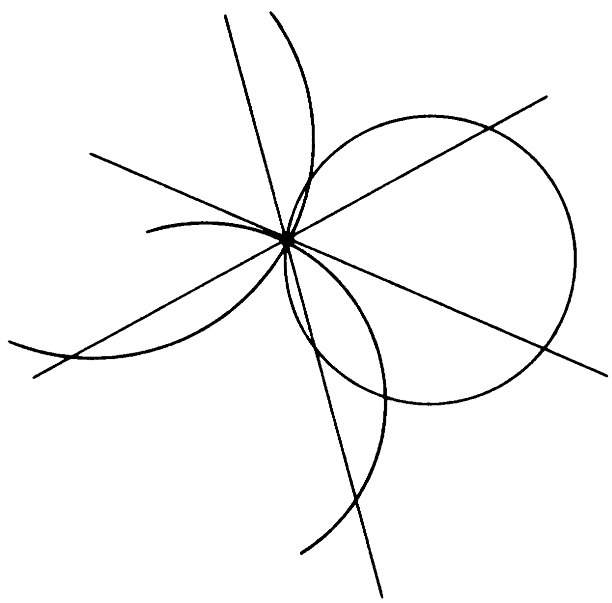
$$C = 0 \quad \text{or} \quad Aw\bar{w} + Bw - \bar{B}\bar{w} + C = 0$$



a



b



c

FIG. 62C

which form a bundle with center O (or w) and power 0 (a ***parabolic*** or ***degenerate*** bundle; by contrast, a bundle of power $k \neq 0$ is called ***nondegenerate***). It is clear that ***a parabolic bundle with center*** O (or w) ***consists of all circles*** (and lines) ***passing through*** O (or w), since the number 0 (or w) satisfies the equations of all circles of the bundle. It is clear that any circular transformation takes a parabolic bundle into

a parabolic bundle (Figure 62C,c); however, this statement is completely accurate only if we include among parabolic bundles the set of all lines (*singular parabolic bundle*; in fact, a circular transformation can take the center of a parabolic bundle into the point at infinity ∞, and the set of all "circles through ∞" is the set of lines*).

Finally, a bundle (Equation 19a, or Equation 19 with condition 19′a) whose power k is negative is called *elliptic*. Among the circles of an elliptic bundle (Equation 19a) belongs the circle Σ with center O and radius $\sqrt{-k}$ (since the equation

$$z\bar{z} + k = 0 \tag{19c}$$

of this circle has the form of Equation 19a). Every circle (Equation 19a) intersects the circle Σ in two points z_1 and z_2, which satisfy the system of equations 19a and 19c, that is, they satisfy Equation 19a and the equation

$$Bz - \bar{B}\bar{z} = 0$$

in other words, it intersects Σ in the same points as the diameter $Bz - \bar{B}\bar{z} = 0$ of the circle Σ. Hence it follows that *an elliptic bundle consists of all circles* (and lines) *which intersect a fixed circle Σ in diametrically opposite points* (Figure 62C,b). We may show that every circular transformation takes an elliptic bundle into an elliptic bundle.

We note that the idea of a bundle can be used for the following (geometrical, or more precisely, "geometrized") definition of the inversion (Equation 16): *An inversion with center O and power k* ($k \neq 0$) *takes each point z of the plane into a second point z', where all circles* (*and lines*) *of the bundle with center O and power k which pass through z meet* (cf. Figure 56a on p. 141).† If we suppose that a (nondegenerate, that is, hyperbolic or elliptic) bundle of circles can also be singular, then this definition includes symmetries about lines as well as inversions (Figure 56b).†

¶J, p. 162. The analogy between the role of the tangential distance of two circles in questions connected with axial circular transformations and the role of the angle between circles in the questions dealt with in Sections 13 and 14 leaps to the eye;

---

* See p. 34.

† It is possible that all circles of the bundle which pass through a fixed point z *touch* each other at this point (this will be the case if the point z lies on the circle Σ of a hyperbolic bundle or on the axis of a singular hyperbolic bundle—see [¶H]). In this case we agree to regard the transformation as taking the point z into itself.

here we shall dwell on it in more detail. In Sections 13 and 14 we considered a circle as a set (locus) of points; intersecting circles S_1 and S_2 (only such circles make an angle) are circles having a common point z (Figure 64A, a). Now let z_1 and z_2 be points of S_1 and S_2 close to z; when these points tend to the point z,

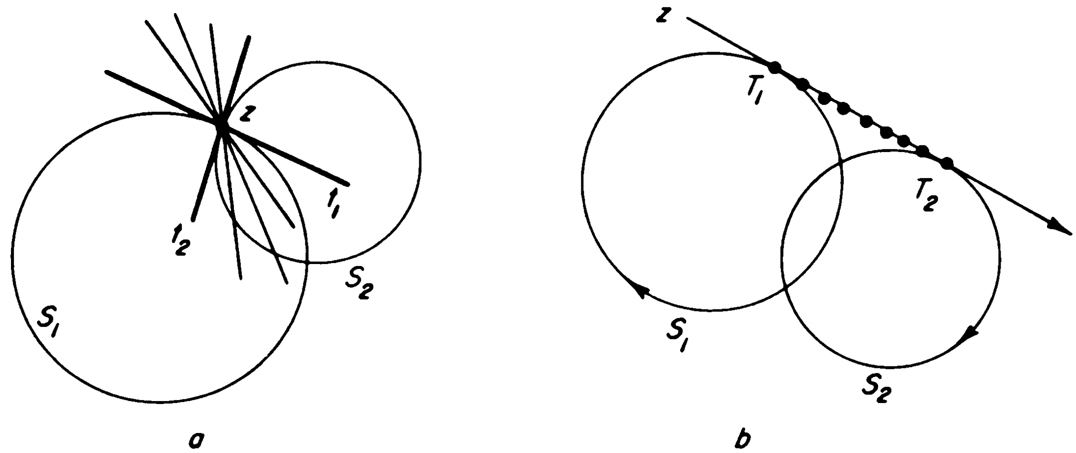


FIG. 64A

the lines $[zz_1]$ and $[zz_2]$ tend to the tangents t_1 and t_2 to S_1 and S_2 at z—lines through z defined at this point by the circles S_1 and S_2 (lines "belonging to" S_1 and S_2). The measure of the set of lines passing through z and included between the lines t_1 and t_2 (see Figure 64A, a) is the angle between t_1 and t_2. This angle is a definite characteristic of the pair of circles S_1 and S_2; it is called the *angle between the circles* S_1 and S_2. In this section we shall consider a circle as a set ("locus") of (oriented) lines; the tangential distance is defined only for two circles S_1 and S_2 which have a common line (common tangent) z (Figure 64A, b). Let z_1 and z_2 be lines of S_1 and S_2 close to the line z; when these lines tend to the line z, the points of intersection of z and z_1, z and z_2, tend to the points T_1 and T_2, defined on the line z by the circles S_1 and S_2 (the points of z "belonging to" S_1 and S_2). The measure of the set of points of the line z between T_1 and T_2—the length of the segment T_1T_2—is an important characteristic of the pair of circles S_1 and S_2; it is called the *tangential distance of these circles*.

¶K, p. 166. This result may be called the ***basic theorem*** *of the theory of axial circular transformations*. Since any motion is the product of not more than three symmetries about lines,* and the symmetry $z' = 1/\bar{z}$ can be regarded as a particular case of an axial inversion $z' = k/\bar{z}$ (an axial inversion of power 1; similarly the reorientation $z' = -1/\bar{z}$ can be called an axial inversion of power -1), it follows from the basic theorem that *every axial circular transformation of the plane can be represented as the product of not more than five axial inversions* (among which are included symmetries about lines and reorientations) *and dilatations.*† We shall give a uniform geometrical description of axial inversions (including here a symmetry about a line) and dilatations in the next section (see p. 237).

¶L, p. 171. It is clear that one of the basic ideas of axial circular geometry is a "circle or point" (a circle of nonzero or zero radius); however, the idea of a point in this geometry is lacking, since an axial circular transformation can take a point into a circle (and a circle into a point). By contrast, the idea of an (oriented) line plays an independent role in axial circular geometry, since an axial circular transformation is defined as a transformation in the set of directed lines (axes); it takes each axis into an axis. The idea of the tangential distance of two circles plays an important role in axial circular geometry (see p. 161); also, the idea of touching circles is significant (since an axial circular transformation takes touching circles‡ into touching circles).

---

* See, for example, I. M. Yaglom, *Geometrical Transformations* (Random House, New York), 1962.

† Three "axial inversions" (in fact, symmetries about lines) may be required for a motion; in the domain of oriented lines, a motion obtained as a product of these three symmetries may differ from the one we need in respect of the directions of the lines, and we may still need a reorientation (a fourth "axial inversion"); thus in all we obtain $3 + 1 + 1 = 5$ (or less) axial inversions and dilatations.

‡ See footnote 81 on p. 162.

¶M, p. 178. An axial inversion (Equation 25b) takes an arbitrary circle or point (Equation 19) into a circle or point

$$Cz\bar{z} + Bkz - \bar{B}k\bar{z} + Ak^2 = 0 \tag{19b}$$

which is generally distinct from the original circle. If Equations 19 and 19b give two different circles S and S' (Figure 71A), then

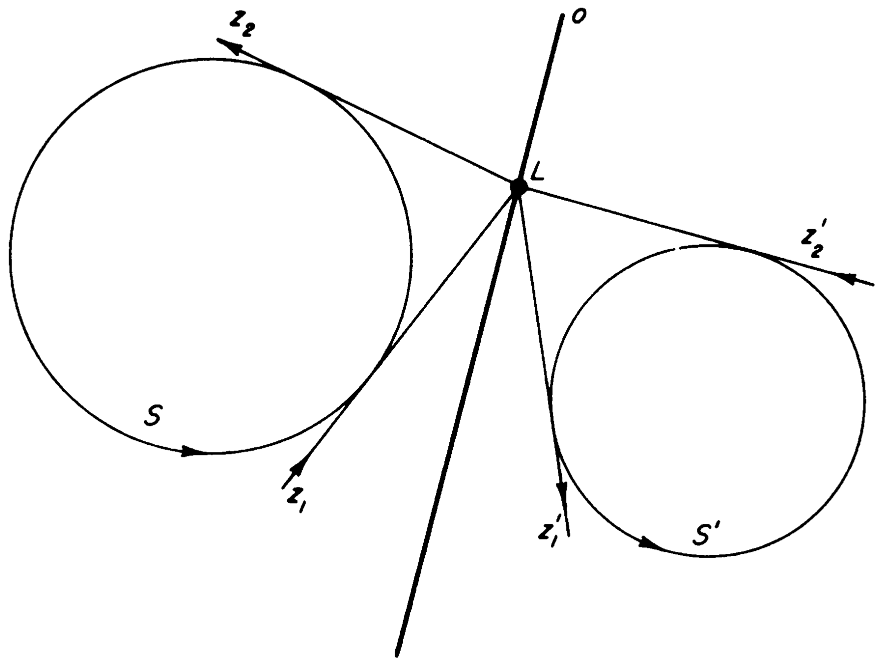


FIG. 71A

since each point L of the line o is taken into itself by an axial inversion with axis o (this follows immediately from the geometrical description of axial inversion), the inversion (Equation 25b) takes the pair of "circles" S and L into the pair of "circles" S' and L. And since every axial circular transformation preserves tangential distances, the "tangential distance" between L and S (the length of the segment of the tangent drawn from L to S) is equal to the "tangential distance" between L and S' (the length of the segment of the tangent drawn from L to S'). [More precisely, if it is possible to draw two tangents z_1 and z_2 from the point L to the circle S, and the inversion (Equation 25b) takes them into tangents z_1' and z_2' of the circle S', then $\{Lz_1S\} = -\{Lz_1'S'\}$, and so $\{Lz_1S\} = \{Lz_2'S'\}$.] Thus *the*

axis o of the inversion (Equation 25b) *which takes S into S' is the radical axis* (see p. 47) *of S and S'.*

This result has an interesting application. Let S be an arbitrary (oriented) circle, considered as the set of its (oriented) tangents, and z_1, z_2, z_3, z_4, z_5, z_6 six arbitrary tangents of this circle, forming a hexagon $\overline{z_1z_2z_3z_4z_5z_6}$ circumscribed to S (Figure 71B;

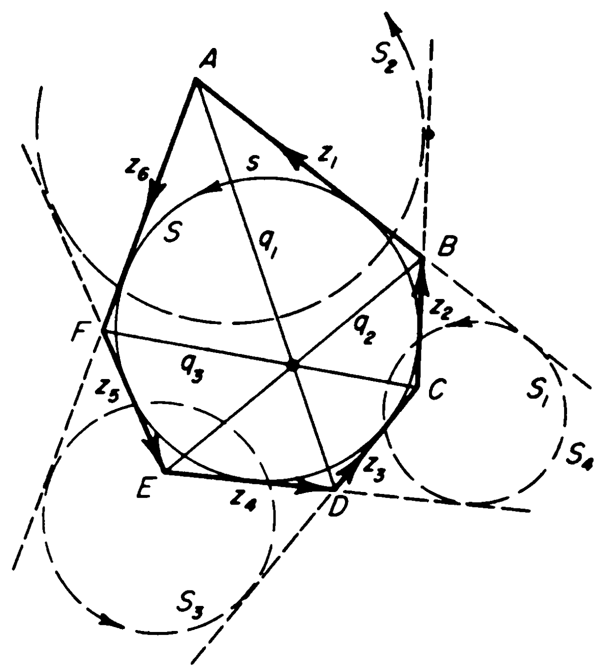


FIG. 71B

this hexagon can be self-intersecting). We denote by A, B, C, D, E, F the vertices of the hexagon $\overline{z_1z_2z_3z_4z_5z_6}$, the points of intersection of z_1 and z_2, z_2 and z_3, z_3 and z_4, z_4 and z_5, z_5 and z_6, z_6 and z_1; we denote the diagonals AD, BE, and CF of the hexagon, joining opposite vertices, by q_1, q_2, and q_3. An axial inversion with axis q_1 and director circle S obviously takes z_1 and z_4 into z_2 and z_5, and vice versa; hence it takes an arbitrary circle S_1 touching the (oriented) lines z_1 and z_4 into a circle S_2 touching z_2 and z_5. Similarly an axial inversion with axis q_2 and director circle S takes z_2 and z_5 into z_3 and z_6, and vice versa; hence this inversion takes the circle S_2 into a circle S_3 touching the lines z_3 and z_6. Finally, a third axial inversion with axis q_3 and the same director circle S takes z_3 and z_6 into z_4 and z_1 and vice versa;

this inversion takes S_3 into a circle S_4 touching the same lines z_4 and z_1 as S_1 does. And since all these axial inversions take S into itself (since S is the director circle of all these inversions), and any axial inversion preserves the absolute value of the tangential distance between two circles and changes the sign of this distance, we have

$$\{S_1z_1S\} = -\{S_2z_2S\} = \{S_3z_3S\} = -\{S_4z_4S\} = \{S_4z_1S\}$$

(the last equation is derived from the fact that the tangential distances of two fixed circles, measured along two different tangents, have opposite signs, as is easily seen; cf. Figure 36, a and b). But from the equation

$$\{S_1z_1S\} = \{S_4z_1S\}$$

it follows that the circles S_1 and S_4 coincide (since there is *only one* circle S_1 which touches two given tangents z_1 and z_4 of a given circle S and such that the tangential distance $\{S_zS_1\} = a$ of S_1 and S has a definite value and a definite sign).

Thus we conclude that the lines q_1, q_2, and q_3 are the axes of three axial inversions which take S_1 into S_2, S_2 into S_3, and finally S_3 into S_1. Thus the lines q_1, q_2, and q_3 are the *radical axes* of pairs of the circles S_1, S_2, and S_3. But we know that the three radical axes of pairs of three circles meet in one point or are parallel (see p. 48). Thus we conclude that *if ABCDEF is an arbitrary hexagon circumscribed to a circle S, then the diagonals, joining opposite vertices of the hexagon, meet in a point or are parallel.* This theorem (more precisely, a somewhat more general theorem) was first proved in 1806 by the French geometer C. Brianchon (1760–1854); it is therefore called *Brianchon's theorem.*

We now return to the idea of a net of circles. It is clear that *a net of circles can be defined as the set of all circles with respect to which a fixed line o (or w) has a given power k*; the line is called the axis of the net, and the number k the power of the net. Hence it follows that the circles of a net with axis o and power k are described by equations of the form 19a,

and the circles of a net with arbitrary axis w and power k by equations of the form 19 such that

$$(A - kC)w\bar{w} + (1 + k)Bw - (1 + k)\bar{B}\bar{w} + (C - kA) = 0$$

(cf. Equation 47, p. 103). Hence we may deduce purely analytically that *every axial circular transformation* (Equation 1 or 1a) *takes every net of circles into a net of circles* (that is, the idea of a net of circles *belongs to axial circular geometry*); however, the presence of the singular numbers $c\omega$ and ∞ in the set of dual numbers leads to the necessity of extending the idea of a net of circles by introducing the so-called singular nets, which we shall mention below. [We note that a (nonsingular) net of circles can also be defined as the set of all circles such that *any three of them have the same axis of similitude* (which coincides with the axis of the net).]

We sometimes call a net of power $k > 0$ a *hyperbolic* net, a net of power 0 a *parabolic* or *degenerate* net (in contrast to this, all nets of power $k \neq 0$ are called *nondegenerate*), and finally a net of power $k < 0$ an *elliptic* net.* We may show that every axial circular transformation takes a hyperbolic (or parabolic or elliptic) net of circles into a hyperbolic (or parabolic or elliptic) net; however, in order that this statement should be quite precise, we should adjoint to elliptic nets any set of circles of fixed (positive or nonpositive) radius r (a *singular elliptic net*). In particular, if $r = 0$ we obtain the set of all *points* of the plane, which can therefore be called a singular "net of circles."

We note that the idea of a net of circles can be used for the following description of an axial inversion (Equation 25b): An inversion with axis o and power k ($k \neq 0$) takes each (oriented) line z of the plane into a line z' such that all circles of the net with axis o and power k which touch z also touch z' (cf. the transform of the line z_0 in Figure 71). If we assume that the net (understood to be nondegenerate) of circles which appears in this definition can also be singular, that is, it can be a net of circles of fixed radius r, then this definition includes, as well as an axial inversion, also an *axial dilatation* of magnitude r (see Figure 71C). [It is not difficult to see that if $r = 0$ our definition leads to a reorientation and if $k = 1$ to the symmetry about the line o.]

We shall find that the remarkable analogy between the contents of Sections 13–14 and Sections 15–16 leads to the idea of the existence of a more general theory, special cases of which are the theory of Möbius circular transformations and Laguerre axial circular transformations. This is in fact so, and we shall very briefly sketch the outlines of this more general theory. Let Σ and S be two arbitrary (oriented) circles, among which we now include points (these have no orientation) and

---

* See p. 176, lines 8–15, and particularly Figure 70 on p. 177.

lines; we merely stipulate that Σ is not a point. In this case there exists a circular point transformation which takes Σ into a line Σ_1 and the

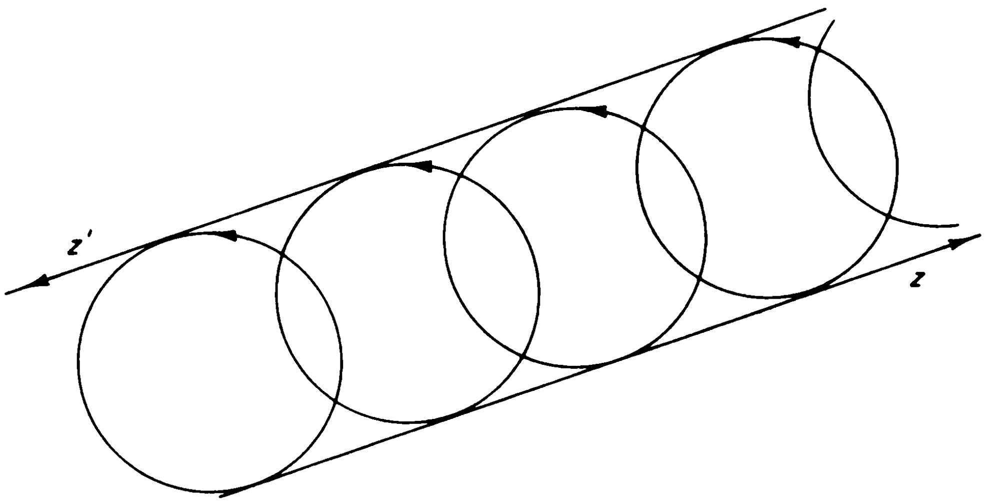


FIG. 71C

circle S into a new circle (not a line) S_1. The power of Σ_1 with respect to S_1 (in the sense of Section 10) is called the **power of the circle Σ with respect to the circle** S; we may show that it does not depend on the choice of the circular transformation which appears in our definition. [This definition is equivalent to the following: Let an arbitrary circle σ, orthogonal to Σ and S, intersect the circle Σ at points A and B and the circle S at points C and D (Figure 71D); in this case the cross-ratio of the distances between the points A, B, C, and D (in the sense of Section 13), taken with a minus sign, is equal to the power of Σ with respect to S.] The set of all circles of the plane, with respect to which the power of a fixed circle Σ is equal to a fixed number k is called a **system of circles** with *central circle Σ and power k*; if the circle Σ is in fact a line, then the system of circles reduces to a *net* of circles, and if $k = 1$ it reduces to a (*hyperbolic*) *bundle* of circles. All the circles of the system which touch a fixed circle S also touch a circle S'; we say that the circle S is taken into the circle S' by a **contact inversion**, whose *central circle* is the central circle of the system and whose *power* is equal to the power of the system. We may show that every contact inversion takes two touching circles into two touching circles. If the central circle of a contact inversion reduces to a line, then the transformation takes each line (circle of infinite radius) into a line; such a "contact inversion" reduces to an axial inversion. If the power k of a contact inversion is equal to 1, then our transformation takes any point (circle of zero radius) into a point and coincides with some (point) inversion (in the sense of Section 13). The set of all transformations generated by all possible contact inversions coincides with all those transformations in the set of (directed) circles,

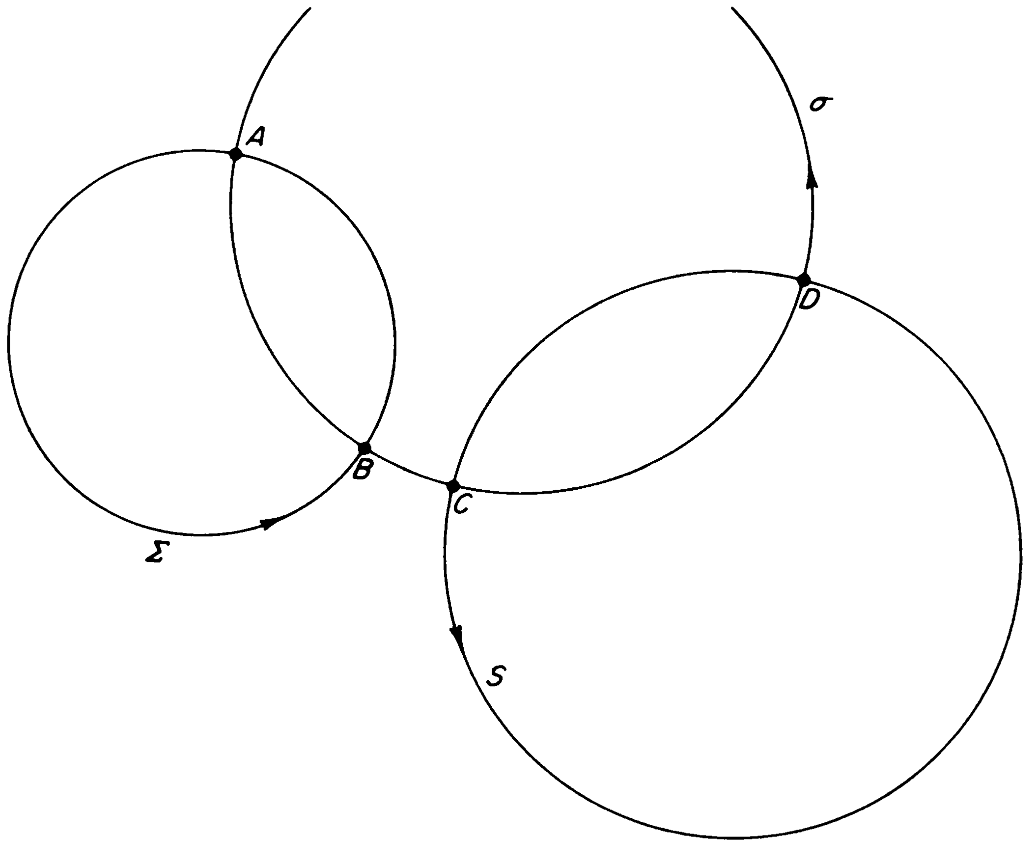


FIG. 71D

which *take touching circles into touching circles*; here a point and a line do not play an independent role, and are considered as particular cases of circles. All such transformations are called **circular contact transformations**; since they were first considered by the outstanding Norwegian mathematician S. Lie (1842–1898), these transformations are also called *Lie (circular) transformations*. Those properties of geometrical figures which are preserved by all circular contact transformations form the subject of **circular contact geometry**; however, we cannot go into the content of this interesting branch of geometry.

This page intentionally left blank

Index